

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 12, 2003, 16:20:05 ; Search time 45 Seconds
(without alignments)
1368.576 Million cell updates/sec

Title: US-10-024-130A-6

Perfect score: 2051

Sequence: 1 MSLGTMRLYPDDIYPLLMK.....CRDAGVLQNRKSYVNDKGP 388

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_19Jun03.*

```

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|---------------------|
| 1 | 2051 | 100.0 | 410 | 21 | Arabidopsis thalia |
| 2 | 2051 | 100.0 | 410 | 21 | Arabidopsis thalia |
| 3 | 2021 | 98.5 | 404 | 21 | Arabidopsis thalia |
| 4 | 2021 | 98.5 | 404 | 21 | Arabidopsis thalia |
| 5 | 1950 | 95.1 | 391 | 21 | Arabidopsis thalia |
| 6 | 1950 | 95.1 | 391 | 21 | Arabidopsis thalia |
| 7 | 1618 | 78.9 | 411 | 17 | Squalene synthetase |
| 8 | 1458 | 71.1 | 395 | 21 | Arabidopsis thalia |
| 9 | 1417 | 69.1 | 392 | 21 | Arabidopsis thalia |

| | | | | | | |
|----|-------|------|-----|----|-----------|---------------------|
| 10 | 1417 | 69.1 | 387 | 21 | AAG277739 | Arabidopsis thalia |
| 11 | 1274 | 62.1 | 264 | 21 | AAG37630 | Arabidopsis thalia |
| 12 | 1274 | 62.1 | 359 | 21 | AAG27740 | Arabidopsis thalia |
| 13 | 1253 | 61.1 | 260 | 21 | AAG37631 | Arabidopsis thalia |
| 14 | 848 | 41.3 | 416 | 23 | AB57061 | Mouse ischaemic co |
| 15 | 820 | 40.0 | 417 | 15 | AA52506 | Human squalene syn |
| 16 | 820 | 40.0 | 417 | 18 | AAW01739 | Human squalene syn |
| 17 | 735 | 35.8 | 448 | 23 | ABP73518 | Candida albicans e |
| 18 | 723 | 35.3 | 444 | 13 | AA222601 | Squalene synthetase |
| 19 | 722 | 35.2 | 444 | 18 | AAW01737 | S. cerevisiae squa |
| 20 | 716.5 | 34.9 | 374 | 24 | ABG72689 | Human squalene syn |
| 21 | 715 | 34.9 | 162 | 21 | AAG23993 | Arabidopsis thalia |
| 22 | 709.5 | 34.6 | 441 | 23 | AAU11290 | Ashbya gossypii AG |
| 23 | 709 | 34.6 | 159 | 21 | AA18097 | Eucalyptus grandis |
| 24 | 700.5 | 34.2 | 460 | 18 | AAW01738 | S. pombe squalene |
| 25 | 675.5 | 32.9 | 520 | 24 | ABJ25847 | Aspergillus fumiga |
| 26 | 675.5 | 32.9 | 528 | 24 | ABJ26447 | Aspergillus fumiga |
| 27 | 582 | 28.4 | 134 | 21 | AAG23994 | Arabidopsis thalia |
| 28 | 561 | 27.4 | 130 | 21 | AAG23995 | Arabidopsis thalia |
| 29 | 480 | 23.4 | 117 | 21 | AA18080 | Pinus radiata squa |
| 30 | 239 | 11.7 | 129 | 21 | AAG01369 | Human secreted pro |
| 31 | 131 | 6.4 | 279 | 23 | ABP53576 | Human NOV8 protein |
| 32 | 131 | 6.4 | 314 | 24 | AAE31589 | Rhodococcus erythr |
| 33 | 128.5 | 6.3 | 290 | 24 | ABP79299 | N. gonorrhoeae ami |
| 34 | 123.5 | 6.0 | 429 | 19 | AAW46964 | Amino acid sequenc |
| 35 | 118 | 5.8 | 425 | 23 | AAG78479 | Maize phytoene syn |
| 36 | 116 | 5.7 | 296 | 23 | AAE22315 | Pantoea stewartii |
| 37 | 116 | 5.7 | 296 | 24 | ABP96689 | Pantoea stewartii |
| 38 | 116 | 5.7 | 296 | 24 | AAQ16022 | Pantoea stewartii |
| 39 | 116 | 5.7 | 425 | 19 | AAW46962 | Amino acid sequenc |
| 40 | 115 | 5.6 | 296 | 20 | AAW90997 | Erwinia uredovora |
| 41 | 115 | 5.6 | 296 | 20 | AAW87887 | Protein encoded by |
| 42 | 115 | 5.6 | 309 | 18 | AAW93845 | E. uredovora phyto |
| 43 | 115 | 5.6 | 309 | 19 | AAW82256 | C. utilis crtB pro |
| 44 | 114.5 | 5.6 | 296 | 11 | AAE07467 | Polypeptide with e |
| 45 | 112.5 | 5.5 | 428 | 19 | AAW46963 | Amino acid sequenc |

ALIGNMENTS

RESULT 1
AAG16425
ID AAG16425 standard; Protein; 410 AA.

XX AC AAG16425;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 17068.

XX DE Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; Genetic mapping; Gene expression control; promoter;
XX KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

XX PR 05-MAR-1999; 99US-0123180.

XX PR 23-MAR-1999; 99US-0123548.

XX PR 25-MAR-1999; 99US-0125788.

XX PR 29-MAR-1999; 99US-0126264.

XX PR 01-APR-1999; 99US-0127462.

XX PR 06-APR-1999; 99US-0128234.

XX PR 08-APR-1999; 99US-0128714.

XX PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-APR-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.

| | | | | | |
|----|--------------|--|-----|--------------|--|
| PR | 01-JUL-1999; | 99US-0141842. | PR | 16-SEP-1999; | 99US-0154039. |
| PR | 01-JUL-1999; | 99US-0142154. | PR | 20-SEP-1999; | 99US-0154779. |
| PR | 02-JUL-1999; | 99US-0142055. | PR | 22-SEP-1999; | 99US-0155139. |
| PR | 06-JUL-1999; | 99US-0142390. | PR | 23-SEP-1999; | 99US-0155486. |
| PR | 08-JUL-1999; | 99US-0142803. | PR | 24-SEP-1999; | 99US-0155659. |
| PR | 09-JUL-1999; | 99US-0142920. | PR | 28-SEP-1999; | 99US-0156458. |
| PR | 12-JUL-1999; | 99US-0142977. | PR | 29-SEP-1999; | 99US-0156596. |
| PR | 13-JUL-1999; | 99US-0143542. | PR | 04-OCT-1999; | 99US-0157117. |
| PR | 14-JUL-1999; | 99US-0143624. | PR | 05-OCT-1999; | 99US-0157753. |
| PR | 15-JUL-1999; | 99US-0144005. | PR | 08-OCT-1999; | 99US-0157865. |
| PR | 16-JUL-1999; | 99US-0144085. | PR | 08-OCT-1999; | 99US-0158029. |
| PR | 16-JUL-1999; | 99US-0144086. | PR | 08-OCT-1999; | 99US-0158232. |
| PR | 19-JUL-1999; | 99US-0144325. | PR | 12-OCT-1999; | 99US-0158369. |
| PR | 19-JUL-1999; | 99US-0144331. | PR | 13-OCT-1999; | 99US-0159293. |
| PR | 19-JUL-1999; | 99US-0144332. | PR | 13-OCT-1999; | 99US-0159294. |
| PR | 19-JUL-1999; | 99US-0144333. | PR | 13-OCT-1999; | 99US-0159295. |
| PR | 19-JUL-1999; | 99US-0144334. | PR | 14-OCT-1999; | 99US-0159329. |
| PR | 19-JUL-1999; | 99US-0144335. | PR | 14-OCT-1999; | 99US-0159330. |
| PR | 20-JUL-1999; | 99US-0144352. | PR | 14-OCT-1999; | 99US-0159331. |
| PR | 20-JUL-1999; | 99US-0144632. | PR | 14-OCT-1999; | 99US-0159637. |
| PR | 20-JUL-1999; | 99US-0144632. | PR | 18-OCT-1999; | 99US-0159638. |
| PR | 21-JUL-1999; | 99US-0144884. | PR | 18-OCT-1999; | 99US-0159584. |
| PR | 21-JUL-1999; | 99US-0145086. | PR | 21-OCT-1999; | 99US-0160741. |
| PR | 21-JUL-1999; | 99US-0145088. | PR | 21-OCT-1999; | 99US-0160767. |
| PR | 21-JUL-1999; | 99US-0145085. | PR | 21-OCT-1999; | 99US-0160768. |
| PR | 22-JUL-1999; | 99US-0145087. | PR | 21-OCT-1999; | 99US-0160770. |
| PR | 22-JUL-1999; | 99US-0145089. | PR | 21-OCT-1999; | 99US-0160814. |
| PR | 22-JUL-1999; | 99US-0145192. | PR | 21-OCT-1999; | 99US-0160815. |
| PR | 23-JUL-1999; | 99US-0145145. | PR | 22-OCT-1999; | 99US-0160815. |
| PR | 23-JUL-1999; | 99US-0145218. | PR | 22-OCT-1999; | 99US-0160980. |
| PR | 23-JUL-1999; | 99US-0145224. | PR | 22-OCT-1999; | 99US-0160981. |
| PR | 26-JUL-1999; | 99US-0145276. | PR | 22-OCT-1999; | 99US-0160989. |
| PR | 27-JUL-1999; | 99US-0145313. | PR | 22-OCT-1999; | 99US-0161404. |
| PR | 27-JUL-1999; | 99US-0145319. | PR | 25-OCT-1999; | 99US-0161405. |
| PR | 27-JUL-1999; | 99US-0145918. | PR | 25-OCT-1999; | 99US-0161406. |
| PR | 28-JUL-1999; | 99US-0145951. | PR | 26-OCT-1999; | 99US-0161359. |
| PR | 02-AUG-1999; | 99US-0146386. | PR | 26-OCT-1999; | 99US-0161360. |
| PR | 02-AUG-1999; | 99US-0146388. | PR | 26-OCT-1999; | 99US-0161361. |
| PR | 02-AUG-1999; | 99US-0146389. | PR | 26-OCT-1999; | 99US-0161920. |
| PR | 03-AUG-1999; | 99US-0147038. | PR | 28-OCT-1999; | 99US-0161992. |
| PR | 04-AUG-1999; | 99US-0147204. | PR | 28-OCT-1999; | 99US-0161993. |
| PR | 04-AUG-1999; | 99US-0147302. | PR | 29-OCT-1999; | 99US-0162142. |
| PR | 05-AUG-1999; | 99US-0147192. | | | |
| PR | 05-AUG-1999; | 99US-0147260. | | | |
| PR | 06-AUG-1999; | 99US-0147303. | | | |
| PR | 06-AUG-1999; | 99US-0147416. | | | |
| PR | 08-AUG-1999; | 99US-0147493. | | | |
| PR | 09-AUG-1999; | 99US-0147935. | | | |
| PR | 10-AUG-1999; | 99US-0148171. | | | |
| PR | 11-AUG-1999; | 99US-0148319. | | | |
| PR | 12-AUG-1999; | 99US-0148341. | | | |
| PR | 13-AUG-1999; | 99US-0148565. | | | |
| PR | 13-AUG-1999; | 99US-0148684. | | | |
| PR | 16-AUG-1999; | 99US-0149368. | | | |
| PR | 17-AUG-1999; | 99US-0149175. | | | |
| PR | 18-AUG-1999; | 99US-0149426. | | | |
| PR | 20-AUG-1999; | 99US-0149722. | | | |
| PR | 20-AUG-1999; | 99US-0149723. | | | |
| PR | 20-AUG-1999; | 99US-0149929. | | | |
| PR | 23-AUG-1999; | 99US-0149902. | | | |
| PR | 23-AUG-1999; | 99US-0149930. | | | |
| PR | 25-AUG-1999; | 99US-0150566. | | | |
| PR | 26-AUG-1999; | 99US-0150884. | | | |
| PR | 27-AUG-1999; | 99US-0151065. | | | |
| PR | 27-AUG-1999; | 99US-0151066. | | | |
| PR | 27-AUG-1999; | 99US-0151080. | | | |
| PR | 30-AUG-1999; | 99US-0151303. | | | |
| PR | 31-AUG-1999; | 99US-0151438. | | | |
| PR | 01-SEP-1999; | 99US-0151930. | | | |
| PR | 07-SEP-1999; | 99US-0152363. | | | |
| PR | 10-SEP-1999; | 99US-0153070. | | | |
| PR | 13-SEP-1999; | 99US-0153758. | | | |
| PR | 15-SEP-1999; | 99US-0154018. | | | |
| QY | 1 | MGSLGTMRLYPDDIYPLLMKRAIEKAEKQIPPEPHMGFCYSMLHKVSRFSFLVIOQLNT | 60 | 1 | MGSLGTMRLYPDDIYPLLMKRAIEKAEKQIPPEPHMGFCYSMLHKVSRFSFLVIOQLNT |
| Db | 1 | MGSLGTMRLYPDDIYPLLMKRAIEKAEKQIPPEPHMGFCYSMLHKVSRFSFLVIOQLNT | 60 | 1 | MGSLGTMRLYPDDIYPLLMKRAIEKAEKQIPPEPHMGFCYSMLHKVSRFSFLVIOQLNT |
| QY | 61 | ELRNACVFYLVLRALDTVEDDTSIPTDEKVPILIAEHRHIYDTHVSCGTKEYKILMD | 120 | 61 | ELRNACVFYLVLRALDTVEDDTSIPTDEKVPILIAEHRHIYDTHVSCGTKEYKILMD |
| Db | 61 | ELRNACVFYLVLRALDTVEDDTSIPTDEKVPILIAEHRHIYDTHVSCGTKEYKILMD | 120 | 61 | ELRNACVFYLVLRALDTVEDDTSIPTDEKVPILIAEHRHIYDTHVSCGTKEYKILMD |
| QY | 121 | QFHVSAAFLELEKGYQEAIEEITRRMGAGNAKEICQEVETVDDYDEYCHYVAGLVGL | 180 | 121 | QFHVSAAFLELEKGYQEAIEEITRRMGAGNAKEICQEVETVDDYDEYCHYVAGLVGL |
| Db | 121 | QFHVSAAFLELEKGYQEAIEEITRRMGAGNAKEICQEVETVDDYDEYCHYVAGLVGL | 180 | 121 | QFHVSAAFLELEKGYQEAIEEITRRMGAGNAKEICQEVETVDDYDEYCHYVAGLVGL |
| QY | 181 | SKLFLAAGSEVLTDPWEAISNSMGLFLOKTNIIIRDYLEINEIPKSRMFWPRETWKYAD | 240 | 181 | SKLFLAAGSEVLTDPWEAISNSMGLFLOKTNIIIRDYLEINEIPKSRMFWPRETWKYAD |
| Db | 181 | SKLFLAAGSEVLTDPWEAISNSMGLFLOKTNIIIRDYLEINEIPKSRMFWPRETWKYAD | 240 | 181 | SKLFLAAGSEVLTDPWEAISNSMGLFLOKTNIIIRDYLEINEIPKSRMFWPRETWKYAD |
| QY | 241 | KLEDLYEENTNKSVOCLNEWVTNALMHIEDCLKYMVSLRDPISIFRCAIPIQIMAIGTLA | 300 | 241 | KLEDLYEENTNKSVOCLNEWVTNALMHIEDCLKYMVSLRDPISIFRCAIPIQIMAIGTLA |
| Db | 241 | KLEDLYEENTNKSVOCLNEWVTNALMHIEDCLKYMVSLRDPISIFRCAIPIQIMAIGTLA | 300 | 241 | KLEDLYEENTNKSVOCLNEWVTNALMHIEDCLKYMVSLRDPISIFRCAIPIQIMAIGTLA |
| QY | 301 | LCYNNQVFRGVKLRRLGLTAKVIDRTKTMADYVGFYDFSCMLKTKVDKNDPNASKTLN | 360 | 301 | LCYNNQVFRGVKLRRLGLTAKVIDRTKTMADYVGFYDFSCMLKTKVDKNDPNASKTLN |
| Db | 301 | LCYNNQVFRGVKLRRLGLTAKVIDRTKTMADYVGFYDFSCMLKTKVDKNDPNASKTLN | 360 | 301 | LCYNNQVFRGVKLRRLGLTAKVIDRTKTMADYVGFYDFSCMLKTKVDKNDPNASKTLN |
| QY | 361 | RLEAVQKLCRDAGVLQNRKSYVNDKGQP | 388 | 361 | RLEAVQKLCRDAGVLQNRKSYVNDKGQP |
| Db | 361 | RLEAVQKLCRDAGVLQNRKSYVNDKGQP | 388 | 361 | RLEAVQKLCRDAGVLQNRKSYVNDKGQP |

RESULT 3

AAG16426

ID AAG16426 standard; Protein; 404 AA.

XX AAG16426;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 17069.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EPI033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0132863.

PR 14-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 18-MAY-1999; 99US-0134370.

PR 19-MAY-1999; 99US-0134768.

PR 20-MAY-1999; 99US-0134941.

PR 21-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139763.

PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.

PR 23-JUN-1999; 99US-0140353.

PR 23-JUN-1999; 99US-0140354.

PR 28-JUN-1999; 99US-0140695.

PR 28-JUN-1999; 99US-0140823.

PR 29-JUN-1999; 99US-0140991.

PR 30-JUN-1999; 99US-0141287.

PR 01-JUL-1999; 99US-0141842.

PR 01-JUL-1999; 99US-0142154.

PR 02-JUL-1999; 99US-0142055.

PR 06-JUL-1999; 99US-0142390.

PR 08-JUL-1999; 99US-0142803.

PR 09-JUL-1999; 99US-0142920.

PR 12-JUL-1999; 99US-0142977.

PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.

PR 16-JUL-1999; 99US-0144086.

PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.

PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144884.

PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.

PR 21-JUL-1999; 99US-0145088.

PR 22-JUL-1999; 99US-0145085.

PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145089.

PR 23-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.

PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.

PR 27-JUL-1999; 99US-0145913.

PR 27-JUL-1999; 99US-0145918.

PR 27-JUL-1999; 99US-0145919.

PR 28-JUL-1999; 99US-0145951.

PR 02-AUG-1999; 99US-0146386.

PR 02-AUG-1999; 99US-0146388.

PR 02-AUG-1999; 99US-0146389.

PR 03-AUG-1999; 99US-0147038.

PR 04-AUG-1999; 99US-0147204.

PR 04-AUG-1999; 99US-0147302.

PR 05-AUG-1999; 99US-0147192.

PR 05-AUG-1999; 99US-0147260.

PR 06-AUG-1999; 99US-0147303.

PR 06-AUG-1999; 99US-0147416.

PR 09-AUG-1999; 99US-0147493.

PR 09-AUG-1999; 99US-0147935.

PR 10-AUG-1999; 99US-0148171.

PR 11-AUG-1999; 99US-0148319.

PR 12-AUG-1999; 99US-0148341.

PR 13-AUG-1999; 99US-0148565.

PR 13-AUG-1999; 99US-0148684.

PR 16-AUG-1999; 99US-0149368.

PR 17-AUG-1999; 99US-0149175.

PR 18-AUG-1999; 99US-0149426.

PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142330.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144327.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 08-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 22-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 22-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.

Thu Nov 13 10:01:34 2003

us-10-024-130a-6.rag

```
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match      98.5%; Score 2021; DB 21; Length 404;
Best Local Similarity 100.0%; Pred. No. 8.5e-191;
Matches 382; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 MLRYDDIYPLLMKRAIEKAQKQIPPEPHGFCYSMLHKVSRFSLSVIQQLNTELRNAV 66
   |||||
Db 1 MLRYDDIYPLLMKRAIEKAQKQIPPEPHGFCYSMLHKVSRFSLSVIQQLNTELRNAV 60
   |||||

QY 67 CVFYLVLRALDTVEDDTGIPDEKVPILIAEHRHYDMDHYSCGTKEYKILMDQFHHVS 126
   |||||
Db 61 CVFYLVLRALDTVEDDTGIPDEKVPILIAEHRHYDMDHYSCGTKEYKILMDQFHHVS 120
   |||||

QY 127 AAFLELEKGYQBAIEIITRRMGAGNAKFCQEVETVDDYDEYCHVAGLVGLGSLKFLA 186
   |||||
Db 121 AAFLELEKGYQBAIEIITRRMGAGNAKFCQEVETVDDYDEYCHVAGLVGLGSLKFLA 180
   |||||

QY 187 AGSEVLTDPWEAISNSMGLFLOKTNIIIRDYLEINEIPKSRMFWPRIWGYADKLEDLK 246
   |||||
Db 181 AGSEVLTDPWEAISNSMGLFLOKTNIIIRDYLEINEIPKSRMFWPRIWGYADKLEDLK 240
   |||||

QY 247 YEENTNKSVOCLNEMVTNALMHIEDCLKIMVSLRDPISFRCAIPQIWAIGTLALCYNNE 306
   |||||
Db 241 YEENTNKSVOCLNEMVTNALMHIEDCLKIMVSLRDPISFRCAIPQIWAIGTLALCYNNE 300
   |||||

QY 307 QVFRGVVKKLRGLTAKVIDRTKTMADVGFYDFSCMLKTKVDKNDPNASKTLNRLEAVQ 366
   |||||
Db 301 QVFRGVVKKLRGLTAKVIDRTKTMADVGFYDFSCMLKTKVDKNDPNASKTLNRLEAVQ 360
   |||||

QY 367 KLCRDAGVLQNRKSYVNDKGQP 388
   |||||
Db 361 KLCRDAGVLQNRKSYVNDKGQP 382
   |||||

RESULT 5
AAG16427
ID AAG16427 standard; Protein; 391 AA.
XX
AC AAG16427;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 17070.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 23-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127482.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 28-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140931.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
```


| | | | | | |
|----|--------------|---------------|--|--|--|
| PR | 21-JUL-1999; | 99US-0145086. | PR | 21-OCT-1999; | 99US-0160741. |
| PR | 21-JUL-1999; | 99US-0145088. | PR | 21-OCT-1999; | 99US-0160767. |
| PR | 22-JUL-1999; | 99US-0145085. | PR | 21-OCT-1999; | 99US-0160768. |
| PR | 22-JUL-1999; | 99US-0145087. | PR | 21-OCT-1999; | 99US-0160770. |
| PR | 22-JUL-1999; | 99US-0145089. | PR | 21-OCT-1999; | 99US-0160814. |
| PR | 22-JUL-1999; | 99US-0145192. | PR | 21-OCT-1999; | 99US-0160815. |
| PR | 23-JUL-1999; | 99US-0145145. | PR | 22-OCT-1999; | 99US-0160980. |
| PR | 23-JUL-1999; | 99US-0145218. | PR | 22-OCT-1999; | 99US-0160981. |
| PR | 23-JUL-1999; | 99US-0145224. | PR | 22-OCT-1999; | 99US-0160989. |
| PR | 26-JUL-1999; | 99US-0145276. | PR | 25-OCT-1999; | 99US-0161404. |
| PR | 27-JUL-1999; | 99US-0145913. | PR | 25-OCT-1999; | 99US-0161405. |
| PR | 27-JUL-1999; | 99US-0145918. | PR | 25-OCT-1999; | 99US-0161406. |
| PR | 27-JUL-1999; | 99US-0145919. | PR | 26-OCT-1999; | 99US-0161359. |
| PR | 28-JUL-1999; | 99US-0145951. | PR | 26-OCT-1999; | 99US-0161360. |
| PR | 02-AUG-1999; | 99US-0146386. | PR | 26-OCT-1999; | 99US-0161361. |
| PR | 02-AUG-1999; | 99US-0146388. | PR | 28-OCT-1999; | 99US-0161920. |
| PR | 02-AUG-1999; | 99US-0146389. | PR | 28-OCT-1999; | 99US-0161992. |
| PR | 03-AUG-1999; | 99US-0147038. | PR | 28-OCT-1999; | 99US-0161993. |
| PR | 04-AUG-1999; | 99US-0147204. | PR | 29-OCT-1999; | 99US-0162142. |
| PR | 04-AUG-1999; | 99US-0147302. | Query Match 95.1%; Score 1950; DB 21; Length 391; | | |
| PR | 05-AUG-1999; | 99US-0147192. | Best Local Similarity 100.0%; Pred. No. 8.4e-184; | | |
| PR | 05-AUG-1999; | 99US-0147260. | Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | | |
| PR | 06-AUG-1999; | 99US-0147303. | | | |
| PR | 06-AUG-1999; | 99US-0147416. | | | |
| PR | 09-AUG-1999; | 99US-0147493. | Qy | 20 | MKRAIEKAERKQIPPEPHWGFCSYMLHKVSRFSLSVIQQLNTELRNAVCFYLVLRALDTV 79 |
| PR | 09-AUG-1999; | 99US-0147935. | Db | 1 | MKRAIEKAERKQIPPEPHWGFCSYMLHKVSRFSLSVIQQLNTELRNAVCFYLVLRALDTV 60 |
| PR | 10-AUG-1999; | 99US-0148171. | | | |
| PR | 11-AUG-1999; | 99US-0148319. | Qy | 80 | EDDTSIPTDEKVPILIAFHRHIYDTHWYSCGTKEYKILMDQFHHVSAAFLELEKGYQEA 139 |
| PR | 12-AUG-1999; | 99US-0148341. | Db | 61 | EDDTSIPTDEKVPILIAFHRHIYDTHWYSCGTKEYKILMDQFHHVSAAFLELEKGYQEA 120 |
| PR | 13-AUG-1999; | 99US-0148565. | | | |
| PR | 13-AUG-1999; | 99US-0148684. | Qy | 140 | IEEITRMGAGMAKFIQCEVETVDDYDEYCHYVAGLVGLSKLFLAAGSEVLTDPWEAI 199 |
| PR | 16-AUG-1999; | 99US-0149368. | Db | 121 | IEEITRMGAGMAKFIQCEVETVDDYDEYCHYVAGLVGLSKLFLAAGSEVLTDPWEAI 180 |
| PR | 17-AUG-1999; | 99US-0149175. | | | |
| PR | 18-AUG-1999; | 99US-0149426. | Qy | 200 | SNSMGLFLQKTIIRDYLEDINEIPKSRMPFWEIWKYADKLEDLKYEENTNKSVOCLN 259 |
| PR | 20-AUG-1999; | 99US-0149722. | Db | 181 | SNSMGLFLQKTIIRDYLEDINEIPKSRMPFWEIWKYADKLEDLKYEENTNKSVOCLN 240 |
| PR | 20-AUG-1999; | 99US-0149723. | | | |
| PR | 20-AUG-1999; | 99US-0149929. | Qy | 260 | EMVTNALMHIEDCLKYVMSLRDPSIFRFCAIPQIMAGTILALCYNNEQVFRGVVVKLRGL 319 |
| PR | 23-AUG-1999; | 99US-0149902. | Db | 241 | EMVTNALMHIEDCLKYVMSLRDPSIFRFCAIPQIMAGTILALCYNNEQVFRGVVVKLRGL 300 |
| PR | 23-AUG-1999; | 99US-0149930. | | | |
| PR | 25-AUG-1999; | 99US-0150566. | Qy | 320 | TAKVIDRTKTMADYVGFYDFSCMLKTKVDKNDPNASKTLNRLEAVOKLCRDAGVQNRK 379 |
| PR | 26-AUG-1999; | 99US-0150884. | Db | 301 | TAKVIDRTKTMADYVGFYDFSCMLKTKVDKNDPNASKTLNRLEAVOKLCRDAGVQNRK 360 |
| PR | 27-AUG-1999; | 99US-0151065. | | | |
| PR | 27-AUG-1999; | 99US-0151066. | Qy | 380 | SYVNDKGQP 388 |
| PR | 30-AUG-1999; | 99US-0151303. | Db | 361 | SYVNDKGQP 369 |
| PR | 31-AUG-1999; | 99US-0151438. | | | |
| PR | 01-SEP-1999; | 99US-0151930. | RESULT 6 | | |
| PR | 07-SEP-1999; | 99US-0152363. | AAG32433 | | |
| PR | 10-SEP-1999; | 99US-0153070. | ID | AAG32433 standard; Protein; 391 AA. | |
| PR | 13-SEP-1999; | 99US-0153758. | AC | AAG32433; | |
| PR | 15-SEP-1999; | 99US-0154018. | XX | | |
| PR | 16-SEP-1999; | 99US-0154039. | XX | | |
| PR | 20-SEP-1999; | 99US-0154779. | DT | 17-OCT-2000 (first entry) | |
| PR | 22-SEP-1999; | 99US-0155139. | XX | | |
| PR | 23-SEP-1999; | 99US-0155486. | XX | Arabidopsis thaliana protein fragment SEQ ID NO: 39124. | |
| PR | 24-SEP-1999; | 99US-0155659. | DE | | |
| PR | 28-SEP-1999; | 99US-0156458. | XX | | |
| PR | 29-SEP-1999; | 99US-0156596. | KW | Protein identification; signal transduction pathway; metabolic pathway; | |
| PR | 04-OCT-1999; | 99US-0157117. | KW | hybridisation assay; genetic mapping; gene expression control; promoter; | |
| PR | 05-OCT-1999; | 99US-0157753. | KW | termination sequence. | |
| PR | 06-OCT-1999; | 99US-0157865. | XX | | |
| PR | 07-OCT-1999; | 99US-0158029. | OS | Arabidopsis thaliana. | |
| PR | 08-OCT-1999; | 99US-0158232. | XX | | |
| PR | 12-OCT-1999; | 99US-0158369. | XX | | |
| PR | 13-OCT-1999; | 99US-0159293. | PN | EP1033405-A2. | |
| PR | 13-OCT-1999; | 99US-0159294. | XX | | |
| PR | 13-OCT-1999; | 99US-0159295. | XX | | |
| PR | 14-OCT-1999; | 99US-0159329. | PD | 06-SEP-2000. | |
| PR | 14-OCT-1999; | 99US-0159330. | XX | | |
| PR | 14-OCT-1999; | 99US-0159331. | XX | | |
| PR | 14-OCT-1999; | 99US-0159637. | | | |
| PR | 14-OCT-1999; | 99US-0159638. | | | |
| PR | 18-OCT-1999; | 99US-0159584. | | | |

us-10-024-130a-6.rag

Thu Nov 13 10:01:34 2003

| | | | | |
|----|------------------------------|----|--------------|---------------|
| PR | 25-FEB-2000; 2000EP-0301439. | PR | 09-JUL-1999; | 99US-0142920. |
| XX | | PR | 12-JUL-1999; | 99US-0142977. |
| PR | 25-FEB-1999; | PR | 13-JUL-1999; | 99US-0143542. |
| PR | 05-MAR-1999; | PR | 14-JUL-1999; | 99US-0143624. |
| PR | 09-MAR-1999; | PR | 15-JUL-1999; | 99US-0144005. |
| PR | 23-MAR-1999; | PR | 16-JUL-1999; | 99US-0144085. |
| PR | 25-MAR-1999; | PR | 16-JUL-1999; | 99US-0144086. |
| PR | 29-MAR-1999; | PR | 19-JUL-1999; | 99US-0144325. |
| PR | 01-APR-1999; | PR | 19-JUL-1999; | 99US-0144331. |
| PR | 06-APR-1999; | PR | 19-JUL-1999; | 99US-0144332. |
| PR | 08-APR-1999; | PR | 19-JUL-1999; | 99US-0144333. |
| PR | 16-APR-1999; | PR | 19-JUL-1999; | 99US-0144334. |
| PR | 19-APR-1999; | PR | 19-JUL-1999; | 99US-0144335. |
| PR | 21-APR-1999; | PR | 20-JUL-1999; | 99US-0144352. |
| PR | 23-APR-1999; | PR | 20-JUL-1999; | 99US-0144632. |
| PR | 28-APR-1999; | PR | 20-JUL-1999; | 99US-0144884. |
| PR | 30-APR-1999; | PR | 21-JUL-1999; | 99US-0144814. |
| PR | 04-MAY-1999; | PR | 21-JUL-1999; | 99US-0145086. |
| PR | 05-MAY-1999; | PR | 21-JUL-1999; | 99US-0145088. |
| PR | 06-MAY-1999; | PR | 22-JUL-1999; | 99US-0145085. |
| PR | 06-MAY-1999; | PR | 22-JUL-1999; | 99US-0145087. |
| PR | 07-MAY-1999; | PR | 22-JUL-1999; | 99US-0145089. |
| PR | 11-MAY-1999; | PR | 22-JUL-1999; | 99US-0145192. |
| PR | 14-MAY-1999; | PR | 23-JUL-1999; | 99US-0145145. |
| PR | 14-MAY-1999; | PR | 23-JUL-1999; | 99US-0145218. |
| PR | 14-MAY-1999; | PR | 26-JUL-1999; | 99US-0145224. |
| PR | 14-MAY-1999; | PR | 26-JUL-1999; | 99US-0145276. |
| PR | 18-MAY-1999; | PR | 27-JUL-1999; | 99US-0145913. |
| PR | 18-MAY-1999; | PR | 27-JUL-1999; | 99US-0145918. |
| PR | 19-MAY-1999; | PR | 27-JUL-1999; | 99US-0145919. |
| PR | 20-MAY-1999; | PR | 28-JUL-1999; | 99US-0145951. |
| PR | 21-MAY-1999; | PR | 02-AUG-1999; | 99US-0146386. |
| PR | 24-MAY-1999; | PR | 02-AUG-1999; | 99US-0146388. |
| PR | 25-MAY-1999; | PR | 02-AUG-1999; | 99US-0146389. |
| PR | 27-MAY-1999; | PR | 03-AUG-1999; | 99US-0147038. |
| PR | 28-MAY-1999; | PR | 04-AUG-1999; | 99US-0147204. |
| PR | 01-JUN-1999; | PR | 04-AUG-1999; | 99US-0147302. |
| PR | 03-JUN-1999; | PR | 05-AUG-1999; | 99US-0147192. |
| PR | 04-JUN-1999; | PR | 05-AUG-1999; | 99US-0147260. |
| PR | 07-JUN-1999; | PR | 06-AUG-1999; | 99US-0147303. |
| PR | 08-JUN-1999; | PR | 06-AUG-1999; | 99US-0147416. |
| PR | 10-JUN-1999; | PR | 09-AUG-1999; | 99US-0147493. |
| PR | 10-JUN-1999; | PR | 09-AUG-1999; | 99US-0147935. |
| PR | 13-JUN-1999; | PR | 10-AUG-1999; | 99US-0148171. |
| PR | 16-JUN-1999; | PR | 11-AUG-1999; | 99US-0148319. |
| PR | 16-JUN-1999; | PR | 12-AUG-1999; | 99US-0148341. |
| PR | 17-JUN-1999; | PR | 13-AUG-1999; | 99US-0148565. |
| PR | 18-JUN-1999; | PR | 13-AUG-1999; | 99US-0148684. |
| PR | 18-JUN-1999; | PR | 16-AUG-1999; | 99US-0149368. |
| PR | 18-JUN-1999; | PR | 17-AUG-1999; | 99US-0149175. |
| PR | 18-JUN-1999; | PR | 18-AUG-1999; | 99US-0149426. |
| PR | 18-JUN-1999; | PR | 20-AUG-1999; | 99US-0149722. |
| PR | 18-JUN-1999; | PR | 20-AUG-1999; | 99US-0149723. |
| PR | 18-JUN-1999; | PR | 20-AUG-1999; | 99US-0149929. |
| PR | 18-JUN-1999; | PR | 20-AUG-1999; | 99US-0149902. |
| PR | 18-JUN-1999; | PR | 23-AUG-1999; | 99US-0149930. |
| PR | 18-JUN-1999; | PR | 25-AUG-1999; | 99US-0150566. |
| PR | 18-JUN-1999; | PR | 26-AUG-1999; | 99US-0150884. |
| PR | 18-JUN-1999; | PR | 27-AUG-1999; | 99US-0151065. |
| PR | 22-JUN-1999; | PR | 27-AUG-1999; | 99US-0151066. |
| PR | 23-JUN-1999; | PR | 27-AUG-1999; | 99US-0151080. |
| PR | 23-JUN-1999; | PR | 30-AUG-1999; | 99US-0151303. |
| PR | 24-JUN-1999; | PR | 31-AUG-1999; | 99US-0151436. |
| PR | 28-JUN-1999; | PR | 01-SEP-1999; | 99US-0151930. |
| PR | 29-JUN-1999; | PR | 07-SEP-1999; | 99US-0152363. |
| PR | 30-JUN-1999; | PR | 10-SEP-1999; | 99US-0153070. |
| PR | 01-JUL-1999; | PR | 13-SEP-1999; | 99US-0153758. |
| PR | 01-JUL-1999; | PR | 15-SEP-1999; | 99US-0154018. |
| PR | 02-JUL-1999; | PR | 16-SEP-1999; | 99US-0154039. |
| PR | 06-JUL-1999; | PR | 20-SEP-1999; | 99US-0154779. |
| PR | 08-JUL-1999; | PR | 22-SEP-1999; | 99US-0155139. |
| PR | | PR | 23-SEP-1999; | 99US-0155486. |
| PR | | PR | 24-SEP-1999; | 99US-0155659. |

```

PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0162142.

Query Match 95.1%; Score 1950; DB 21; Length 391;
Best Local Similarity 100.0%; Pred. No. 8.4e-184;
Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 MKRAIEAKAEQIPEPHWGFCYSMLHKVSRFSFSLVIOQLNTELRNAVCFYLVLRALDTV 79
DB 1 MKRAIEAKAEQIPEPHWGFCYSMLHKVSRFSFSLVIOQLNTELRNAVCFYLVLRALDTV 60

QY 80 EDTSIPTDEKVPILIAFHRHIYDTHWYSCGTKEYKILMDOPHHVSAAFLELEKGYQEA 139
DB 61 EDTSIPTDEKVPILIAFHRHIYDTHWYSCGTKEYKILMDOPHHVSAAFLELEKGYQEA 120

QY 140 IEETIRRMGAGMAKFIQOEVTVDYDEYCHYVAGLVGLGSLKFLAAGSEVLTDPWEAI 199
DB 121 IEETIRRMGAGMAKFIQOEVTVDYDEYCHYVAGLVGLGSLKFLAAGSEVLTDPWEAI 180

QY 200 SNSMGLFLOKTIIRDYLEDINEIPKSRMFWPREIMGWKYADKLEDLYEENTNKSVOCLN 259
DB 181 SNSMGLFLOKTIIRDYLEDINEIPKSRMFWPREIMGWKYADKLEDLYEENTNKSVOCLN 240

QY 260 EMTNLMHLEDCLKYVMSLRDPSIFRCAIPQIMAGTALCYNNEQVRGVVVKLRGL 319
DB 241 EMTNLMHLEDCLKYVMSLRDPSIFRCAIPQIMAGTALCYNNEQVRGVVVKLRGL 300

QY 320 TAKVIDRTKTMDVYGFYDFSCMLTKTKVDKQNDPNASKTLNRLAEAVOKLCRDAGVLQNRK 379
DB 301 TAKVIDRTKTMDVYGFYDFSCMLTKTKVDKQNDPNASKTLNRLAEAVOKLCRDAGVLQNRK 360

QY 380 SYVNDKGQP 388
DB 361 SYVNDKGQP 369

RESULT 7
AAR94574
ID AAR94574 standard; Protein; 411 AA.

```

```

XX AAR94574;
AC 31-OCT-1996 (first entry)
DT Squalene synthetase from Nicotiana benthamiana.
DE
XX Isoprenoid; sterol; biosynthesis; sterol; transformation; recombinant;
KW genetic engineering; squalene; cholesterol; cycloartenol; tobacco;
KW obtusifoliosol; phytoene; lycopene; beta-carotene; plant.
XX Nicotiana benthamiana.
OS
XX WO9609393-A1.
PN
XX 28-MAR-1996.
XX
XX 07-SEP-1995; 95WO-US11280.
PF
XX
XX 22-SEP-1994; 94US-0310693.
XX (BIOS-) BIOSOURCE TECHNOLOGIES INC.
XX (RETO) REYNOLDS TECHNOLOGIES INC.
XX Hanley KM, Hellmann GM, Nicolas O;
XX WPI; 1996-188455/19.
XX N-PSDB; AAT14599.
DR
XX
XX New nucleic acids encoding squalene synthetase - used to alter the
PT biosynthetic pathway of sterol(s) and isoprenoid(s) or in the in
PT vitro production of such compounds
XX
XX Claim 2; Page 34-35; 46pp; English.
XX
XX AAR94574 is a squalene synthetase enzyme derived from Nicotiana
CC benthamiana. DNA encoding the enzyme can be engineered to be expressed
CC in such plants as tobacco and tomato. To alter the biosynthesis of
CC sterols and other isoprenoids, e.g. cycloartenol, cholesterol,
CC obtusifoliosol, squalene, phytoene, lycopene and beta-carotene. The
CC squalene synthetase enzyme may also be used for the in vitro synthesis
CC of sterols and isoprenoids.
XX
XX Sequence 411 AA;
Query Match 78.9%; Score 1618; DB 17; Length 411;
Best Local Similarity 77.5%; Pred. No. 5.8e-151;
Matches 296; Conservative 45; Mismatches 39; Indels 2; Gaps 1;

QY 1 MGSILGTMLRYPDDIYPLLLKMKRAIEAKAEQIPEPHWGFCYSMLHKVSRFSFSLVIOQLNT 60
DB 1 MGSILGTMLRYPDDIYPLLLKMKRAIEAKAEQIPEPHWGFCYSMLHKVSRFSFSLVIOQLNT 60

QY 61 ELNANVCVFYLVLRALDTVEDDTSIPTDEKVPILIAFHRHIYDTHWYSCGTKEYKILMD 120
DB 61 ELNANVCVFYLVLRALDTVEDDTSIPTDEKVPILIAFHRHIYDTHWYSCGTKEYKILMD 120

QY 121 QFHVHVAALFLEKGYQEAIEETIRRMGAGMAKFIQOEVTVDYDEYCHYVAGLVGLG 180
DB 121 QFHVHVAALFLEKGYQEAIEETIRRMGAGMAKFIQOEVTVDYDEYCHYVAGLVGLG 180

QY 181 SKPLAAGSEVLTDPWEAINSNMGLFLOKTIIRDYLEDINEIPKSRMFWPREIMGWKYAD 240
DB 181 SKPLAAGSEVLTDPWEAINSNMGLFLOKTIIRDYLEDINEIPKSRMFWPREIMGWKYAD 238

QY 241 KLEDLYEENTNKSVOCLNEMVTNLMHLEDCLKYVMSLRDPSIFRCAIPQIMAGTALA 300
DB 239 KLEDLYEENTNKSVOCLNEMVTNLMHLEDCLKYVMSLRDPSIFRCAIPQIMAGTALA 298

QY 301 LCVNNEQVRGVVVKLRGLTAKVIDRTKTMDVYGFYDFSCMLTKTKVDKQNDPNASKTLN 360
DB 299 MCVNDIEVFRGVVVKLRGLTAKVIDRTKTMDVYGFYDFSCMLTKTKVDKQNDPNASKTLN 358

```

Thu Nov 13 10:01:34 2003

us-10-024-130a-6.rag

| | | | | | | |
|----|-----|------------------------|-----|----|--------------|---------------|
| QY | 361 | RLEAVQKLCRDAGVLONRKSYV | 382 | PR | 18-JUN-1999; | 99US-0139454. |
| | | : : : | | PR | 18-JUN-1999; | 99US-0139455. |
| Db | 359 | RLEVILKTCRDSGTLNKRKSYI | 380 | PR | 18-JUN-1999; | 99US-0139456. |
| | | : : : | | PR | 18-JUN-1999; | 99US-0139457. |
| | | | | PR | 18-JUN-1999; | 99US-0139458. |
| | | | | PR | 18-JUN-1999; | 99US-0139459. |
| | | | | PR | 18-JUN-1999; | 99US-0139460. |
| | | | | PR | 18-JUN-1999; | 99US-0139461. |
| | | | | PR | 18-JUN-1999; | 99US-0139462. |
| | | | | PR | 18-JUN-1999; | 99US-0139463. |
| | | | | PR | 18-JUN-1999; | 99US-0139750. |
| | | | | PR | 18-JUN-1999; | 99US-0139753. |
| | | | | PR | 21-JUN-1999; | 99US-0139817. |
| | | | | PR | 22-JUN-1999; | 99US-0139899. |
| | | | | PR | 23-JUN-1999; | 99US-0140353. |
| | | | | PR | 23-JUN-1999; | 99US-0140354. |
| | | | | PR | 24-JUN-1999; | 99US-0140695. |
| | | | | PR | 24-JUN-1999; | 99US-0140823. |
| | | | | PR | 29-JUN-1999; | 99US-0140991. |
| | | | | PR | 30-JUN-1999; | 99US-0141287. |
| | | | | PR | 01-JUL-1999; | 99US-0141842. |
| | | | | PR | 01-JUL-1999; | 99US-0142154. |
| | | | | PR | 02-JUL-1999; | 99US-0142055. |
| | | | | PR | 06-JUL-1999; | 99US-0142390. |
| | | | | PR | 08-JUL-1999; | 99US-0142803. |
| | | | | PR | 09-JUL-1999; | 99US-0142920. |
| | | | | PR | 12-JUL-1999; | 99US-0142977. |
| | | | | PR | 13-JUL-1999; | 99US-0143542. |
| | | | | PR | 14-JUL-1999; | 99US-0143624. |
| | | | | PR | 15-JUL-1999; | 99US-0144005. |
| | | | | PR | 16-JUL-1999; | 99US-0144085. |
| | | | | PR | 16-JUL-1999; | 99US-0144086. |
| | | | | PR | 19-JUL-1999; | 99US-0144325. |
| | | | | PR | 19-JUL-1999; | 99US-0144331. |
| | | | | PR | 19-JUL-1999; | 99US-0144332. |
| | | | | PR | 19-JUL-1999; | 99US-0144333. |
| | | | | PR | 19-JUL-1999; | 99US-0144334. |
| | | | | PR | 19-JUL-1999; | 99US-0144335. |
| | | | | PR | 20-JUL-1999; | 99US-0144352. |
| | | | | PR | 20-JUL-1999; | 99US-0144632. |
| | | | | PR | 20-JUL-1999; | 99US-0144884. |
| | | | | PR | 21-JUL-1999; | 99US-0144814. |
| | | | | PR | 21-JUL-1999; | 99US-0145086. |
| | | | | PR | 21-JUL-1999; | 99US-0145088. |
| | | | | PR | 22-JUL-1999; | 99US-0145085. |
| | | | | PR | 22-JUL-1999; | 99US-0145087. |
| | | | | PR | 22-JUL-1999; | 99US-0145089. |
| | | | | PR | 22-JUL-1999; | 99US-0145192. |
| | | | | PR | 23-JUL-1999; | 99US-0145145. |
| | | | | PR | 23-JUL-1999; | 99US-0145218. |
| | | | | PR | 23-JUL-1999; | 99US-0145224. |
| | | | | PR | 26-JUL-1999; | 99US-0145276. |
| | | | | PR | 27-JUL-1999; | 99US-0145913. |
| | | | | PR | 27-JUL-1999; | 99US-0145918. |
| | | | | PR | 27-JUL-1999; | 99US-0145919. |
| | | | | PR | 28-JUL-1999; | 99US-0145921. |
| | | | | PR | 02-AUG-1999; | 99US-0146386. |
| | | | | PR | 02-AUG-1999; | 99US-0146388. |
| | | | | PR | 02-AUG-1999; | 99US-0146389. |
| | | | | PR | 03-AUG-1999; | 99US-0147038. |
| | | | | PR | 04-AUG-1999; | 99US-0147204. |
| | | | | PR | 04-AUG-1999; | 99US-0147302. |
| | | | | PR | 05-AUG-1999; | 99US-0147192. |
| | | | | PR | 05-AUG-1999; | 99US-0147260. |
| | | | | PR | 06-AUG-1999; | 99US-0147303. |
| | | | | PR | 06-AUG-1999; | 99US-0147416. |
| | | | | PR | 09-AUG-1999; | 99US-0147493. |
| | | | | PR | 09-AUG-1999; | 99US-0147935. |
| | | | | PR | 10-AUG-1999; | 99US-0148171. |
| | | | | PR | 11-AUG-1999; | 99US-0148319. |
| | | | | PR | 12-AUG-1999; | 99US-0148341. |
| | | | | PR | 13-AUG-1999; | 99US-0148565. |
| | | | | PR | 13-AUG-1999; | 99US-0148684. |

RESULT 8
 AAG27738
 ID AAG27738 standard; Protein; 395 AA.
 XX AAG27738;
 XX
 DT 17-OCT-2000 (first entry)
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 32694.
 XX
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX
 OS Arabidopsis thaliana.
 XX
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-0301439.
 XX
 PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0123788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 04-MAY-1999; 99US-0132407.
 PR 05-MAY-1999; 99US-0132484.
 PR 06-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 07-MAY-1999; 99US-0132487.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 21-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.

| | | | | | | | | | |
|--|--------------|--|-----|-----|--------------------------|------------------------------|---------|-----|-----|
| PR | 16-AUG-1999; | 99US-0149368. | QY | 231 | PREIWGKYADKLEDLKYEENTNKS | VQCLNEMVTNALMHIEDCLKYMVSLRDP | SIFRFC | AI | 290 |
| PR | 17-AUG-1999; | 99US-0149175. | Db | 121 | PREIWGKYADKLEDLKYEENTNKS | VQCLNEMVTNALMHIEDCLKYMVSLRDP | SIFRFC | AI | 180 |
| PR | 18-AUG-1999; | 99US-0149426. | QY | 291 | PQIMAIGTTLALCYNNEQVFRGVV | KLRGLTAKVIDRTKTMDVYGAFYDF | SCMLKTK | VDK | 350 |
| PR | 20-AUG-1999; | 99US-0149722. | Db | 181 | PQIMAIGTTLALCYNNEQVFRGVV | KLRGLTAKVIDRTKTMDVYGAFYDF | SCMLKTK | VDK | 240 |
| PR | 20-AUG-1999; | 99US-0149723. | QY | 351 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 388 | |
| PR | 20-AUG-1999; | 99US-0149902. | Db | 241 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 278 | |
| PR | 23-AUG-1999; | 99US-0149902. | QY | 351 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 388 | |
| PR | 25-AUG-1999; | 99US-0150566. | Db | 241 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 278 | |
| PR | 26-AUG-1999; | 99US-0150884. | QY | 351 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 388 | |
| PR | 27-AUG-1999; | 99US-0151065. | Db | 241 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 278 | |
| PR | 27-AUG-1999; | 99US-0151066. | QY | 351 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 388 | |
| PR | 27-AUG-1999; | 99US-0151080. | Db | 241 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 278 | |
| PR | 30-AUG-1999; | 99US-0151303. | QY | 351 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 388 | |
| PR | 31-AUG-1999; | 99US-0151438. | Db | 241 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 278 | |
| PR | 01-SEP-1999; | 99US-0151930. | QY | 351 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 388 | |
| PR | 07-SEP-1999; | 99US-0152363. | Db | 241 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 278 | |
| PR | 10-SEP-1999; | 99US-0153070. | QY | 351 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 388 | |
| PR | 13-SEP-1999; | 99US-0153758. | Db | 241 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 278 | |
| PR | 15-SEP-1999; | 99US-0154018. | QY | 351 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 388 | |
| PR | 16-SEP-1999; | 99US-0154039. | Db | 241 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 278 | |
| PR | 20-SEP-1999; | 99US-0154779. | QY | 351 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 388 | |
| PR | 22-SEP-1999; | 99US-0155139. | Db | 241 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 278 | |
| PR | 24-SEP-1999; | 99US-0155486. | QY | 351 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 388 | |
| PR | 28-SEP-1999; | 99US-0155659. | Db | 241 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 278 | |
| PR | 28-SEP-1999; | 99US-0156458. | QY | 351 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 388 | |
| PR | 29-SEP-1999; | 99US-0156596. | Db | 241 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 278 | |
| PR | 04-OCT-1999; | 99US-0157117. | QY | 351 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 388 | |
| PR | 05-OCT-1999; | 99US-0157553. | Db | 241 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 278 | |
| PR | 06-OCT-1999; | 99US-0157865. | QY | 351 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 388 | |
| PR | 07-OCT-1999; | 99US-0158029. | Db | 241 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 278 | |
| PR | 08-OCT-1999; | 99US-0158232. | QY | 351 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 388 | |
| PR | 12-OCT-1999; | 99US-0158369. | Db | 241 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 278 | |
| PR | 13-OCT-1999; | 99US-0159293. | QY | 351 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 388 | |
| PR | 13-OCT-1999; | 99US-0159294. | Db | 241 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 278 | |
| PR | 13-OCT-1999; | 99US-0159295. | QY | 351 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 388 | |
| PR | 14-OCT-1999; | 99US-0159329. | Db | 241 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 278 | |
| PR | 14-OCT-1999; | 99US-0159330. | QY | 351 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 388 | |
| PR | 14-OCT-1999; | 99US-0159331. | Db | 241 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 278 | |
| PR | 14-OCT-1999; | 99US-0159637. | QY | 351 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 388 | |
| PR | 18-OCT-1999; | 99US-0159638. | Db | 241 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 278 | |
| PR | 21-OCT-1999; | 99US-0159594. | QY | 351 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 388 | |
| PR | 21-OCT-1999; | 99US-0160741. | Db | 241 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 278 | |
| PR | 21-OCT-1999; | 99US-0160767. | QY | 351 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 388 | |
| PR | 21-OCT-1999; | 99US-0160770. | Db | 241 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 278 | |
| PR | 21-OCT-1999; | 99US-0160814. | QY | 351 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 388 | |
| PR | 21-OCT-1999; | 99US-0160815. | Db | 241 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 278 | |
| PR | 22-OCT-1999; | 99US-0160980. | QY | 351 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 388 | |
| PR | 22-OCT-1999; | 99US-0160981. | Db | 241 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 278 | |
| PR | 22-OCT-1999; | 99US-0160989. | QY | 351 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 388 | |
| PR | 25-OCT-1999; | 99US-0161404. | Db | 241 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 278 | |
| PR | 25-OCT-1999; | 99US-0161405. | QY | 351 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 388 | |
| PR | 25-OCT-1999; | 99US-0161406. | Db | 241 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 278 | |
| PR | 26-OCT-1999; | 99US-0161359. | QY | 351 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 388 | |
| PR | 26-OCT-1999; | 99US-0161360. | Db | 241 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 278 | |
| PR | 26-OCT-1999; | 99US-0161361. | QY | 351 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 388 | |
| PR | 28-OCT-1999; | 99US-0161920. | Db | 241 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 278 | |
| PR | 28-OCT-1999; | 99US-0161922. | QY | 351 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 388 | |
| PR | 28-OCT-1999; | 99US-0161933. | Db | 241 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 278 | |
| PR | 29-OCT-1999; | 99US-0162142. | QY | 351 | NDPNASKTLNRLEAVOKL | CRDAGVLQNRKSYVNDKG | QOP | 388 | |
| Query Match 71.1%; Score 1458; DB 21; Length 395; | | | | | | | | | |
| Best Local Similarity 100.0%; Pred. No. 3 5e-135; | | | | | | | | | |
| Matches 278; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | | | | | | | | | |
| QY | 111 | GTKEYKILMDQFHVSAAFLELEKGYQEAIEEITRRMGAGKAFICQEVETVDDYDEYCH | 170 | | | | | | |
| Db | 1 | GTKEYKILMDQFHVSAAFLELEKGYQEAIEEITRRMGAGKAFICQEVETVDDYDEYCH | 60 | | | | | | |
| QY | 171 | YVAGLVGLGSLKFLAAGSEVLPDWEAISNSMGLFLQKTNIRDYLEDINEIPKSRMF | 230 | | | | | | |
| Db | 61 | YVAGLVGLGSLKFLAAGSEVLPDWEAISNSMGLFLQKTNIRDYLEDINEIPKSRMF | 120 | | | | | | |

us-10-024-130a-6.rag

Thu Nov 13 10:01:34 2003

| | | |
|-----------------------|--------------|--------------------------------|
| PR | 04-JUN-1999; | 99US-0137502. |
| PR | 07-JUN-1999; | 99US-0137724. |
| PR | 08-JUN-1999; | 99US-0138094. |
| PR | 10-JUN-1999; | 99US-0138540. |
| PR | 10-JUN-1999; | 99US-0138847. |
| PR | 14-JUN-1999; | 99US-0139119. |
| PR | 16-JUN-1999; | 99US-0139452. |
| PR | 16-JUN-1999; | 99US-0139453. |
| PR | 17-JUN-1999; | 99US-0139459. |
| PR | 17-JUN-1999; | 99US-0139452. |
| PR | 18-JUN-1999; | 99US-0139454. |
| PR | 18-JUN-1999; | 99US-0139456. |
| PR | 18-JUN-1999; | 99US-0139457. |
| PR | 18-JUN-1999; | 99US-0139458. |
| PR | 18-JUN-1999; | 99US-0139459. |
| PR | 18-JUN-1999; | 99US-0139460. |
| PR | 18-JUN-1999; | 99US-0139461. |
| PR | 18-JUN-1999; | 99US-0139462. |
| PR | 18-JUN-1999; | 99US-0139463. |
| PR | 18-JUN-1999; | 99US-0139750. |
| PR | 18-JUN-1999; | 99US-0139763. |
| PR | 21-JUN-1999; | 99US-0139817. |
| PR | 22-JUN-1999; | 99US-0139899. |
| PR | 23-JUN-1999; | 99US-0140353. |
| PR | 23-JUN-1999; | 99US-0140354. |
| PR | 24-JUN-1999; | 99US-0140695. |
| PR | 28-JUN-1999; | 99US-0140823. |
| PR | 29-JUN-1999; | 99US-0140991. |
| PR | 30-JUN-1999; | 99US-0141287. |
| PR | 01-JUL-1999; | 99US-0141842. |
| PR | 01-JUL-1999; | 99US-0142154. |
| PR | 02-JUL-1999; | 99US-0142055. |
| PR | 06-JUL-1999; | 99US-0142390. |
| PR | 08-JUL-1999; | 99US-0142803. |
| PR | 09-JUL-1999; | 99US-0142920. |
| PR | 12-JUL-1999; | 99US-0142977. |
| PR | 13-JUL-1999; | 99US-0143542. |
| PR | 14-JUL-1999; | 99US-0143624. |
| PR | 16-JUL-1999; | 99US-0144085. |
| PR | 16-JUL-1999; | 99US-0144086. |
| PR | 19-JUL-1999; | 99US-0144325. |
| PR | 19-JUL-1999; | 99US-0144331. |
| PR | 19-JUL-1999; | 99US-0144332. |
| PR | 19-JUL-1999; | 99US-0144333. |
| PR | 19-JUL-1999; | 99US-0144334. |
| PR | 19-JUL-1999; | 99US-0144335. |
| PR | 20-JUL-1999; | 99US-0144352. |
| PR | 20-JUL-1999; | 99US-0144632. |
| PR | 20-JUL-1999; | 99US-0144884. |
| PR | 21-JUL-1999; | 99US-0144814. |
| PR | 21-JUL-1999; | 99US-0145086. |
| PR | 21-JUL-1999; | 99US-0145088. |
| PR | 22-JUL-1999; | 99US-0145085. |
| PR | 22-JUL-1999; | 99US-0145087. |
| PR | 22-JUL-1999; | 99US-0145192. |
| PR | 23-JUL-1999; | 99US-0145145. |
| PR | 23-JUL-1999; | 99US-0145218. |
| PR | 23-JUL-1999; | 99US-0145224. |
| PR | 26-JUL-1999; | 99US-0145276. |
| PR | 27-JUL-1999; | 99US-0145913. |
| PR | 27-JUL-1999; | 99US-0145918. |
| PR | 27-JUL-1999; | 99US-0145919. |
| PR | 28-JUL-1999; | 99US-0145951. |
| PR | 02-AUG-1999; | 99US-0146386. |
| PR | 02-AUG-1999; | 99US-0146388. |
| PR | 02-AUG-1999; | 99US-0146389. |
| PR | 03-AUG-1999; | 99US-0147038. |
| PR | 04-AUG-1999; | 99US-0147204. |
| PR | 04-AUG-1999; | 99US-0147302. |
| PR | 05-AUG-1999; | 99US-0147192. |
| PR | 05-AUG-1999; | 99US-0147260. |
| PR | 06-AUG-1999; | 99US-0147303. |
| PR | 06-AUG-1999; | 99US-0147416. |
| PR | 09-AUG-1999; | 99US-0147493. |
| PR | 09-AUG-1999; | 99US-0147935. |
| PR | 10-AUG-1999; | 99US-0148171. |
| PR | 11-AUG-1999; | 99US-0148319. |
| PR | 12-AUG-1999; | 99US-0148341. |
| PR | 13-AUG-1999; | 99US-0148565. |
| PR | 13-AUG-1999; | 99US-0148684. |
| PR | 16-AUG-1999; | 99US-0149368. |
| PR | 17-AUG-1999; | 99US-0149175. |
| PR | 18-AUG-1999; | 99US-0149426. |
| PR | 20-AUG-1999; | 99US-0149722. |
| PR | 20-AUG-1999; | 99US-0149723. |
| PR | 20-AUG-1999; | 99US-0149929. |
| PR | 23-AUG-1999; | 99US-0149902. |
| PR | 23-AUG-1999; | 99US-0149920. |
| PR | 25-AUG-1999; | 99US-0150566. |
| PR | 26-AUG-1999; | 99US-0150884. |
| PR | 27-AUG-1999; | 99US-0151065. |
| PR | 27-AUG-1999; | 99US-0151066. |
| PR | 30-AUG-1999; | 99US-0151080. |
| PR | 31-AUG-1999; | 99US-0151303. |
| PR | 01-SEP-1999; | 99US-0151438. |
| PR | 07-SEP-1999; | 99US-0151930. |
| PR | 10-SEP-1999; | 99US-0152363. |
| PR | 13-SEP-1999; | 99US-0153070. |
| PR | 13-SEP-1999; | 99US-0153758. |
| PR | 15-SEP-1999; | 99US-0154018. |
| PR | 16-SEP-1999; | 99US-0154039. |
| PR | 20-SEP-1999; | 99US-0154779. |
| PR | 22-SEP-1999; | 99US-0155139. |
| PR | 23-SEP-1999; | 99US-0155486. |
| PR | 28-SEP-1999; | 99US-0155659. |
| PR | 29-SEP-1999; | 99US-0156458. |
| PR | 04-OCT-1999; | 99US-0157117. |
| PR | 06-OCT-1999; | 99US-0157753. |
| PR | 07-OCT-1999; | 99US-0157865. |
| PR | 08-OCT-1999; | 99US-0158029. |
| PR | 12-OCT-1999; | 99US-0158232. |
| PR | 13-OCT-1999; | 99US-0158369. |
| PR | 13-OCT-1999; | 99US-0159293. |
| PR | 13-OCT-1999; | 99US-0159294. |
| PR | 14-OCT-1999; | 99US-0159329. |
| PR | 14-OCT-1999; | 99US-0159330. |
| PR | 14-OCT-1999; | 99US-0159331. |
| PR | 14-OCT-1999; | 99US-0159637. |
| PR | 14-OCT-1999; | 99US-0159638. |
| PR | 18-OCT-1999; | 99US-0159584. |
| PR | 21-OCT-1999; | 99US-0160741. |
| PR | 21-OCT-1999; | 99US-0160767. |
| PR | 21-OCT-1999; | 99US-0160768. |
| PR | 21-OCT-1999; | 99US-0160770. |
| PR | 21-OCT-1999; | 99US-0160814. |
| PR | 21-OCT-1999; | 99US-0160815. |
| PR | 22-OCT-1999; | 99US-0160980. |
| PR | 22-OCT-1999; | 99US-0160981. |
| PR | 22-OCT-1999; | 99US-0160989. |
| PR | 25-OCT-1999; | 99US-0161404. |
| PR | 25-OCT-1999; | 99US-0161405. |
| PR | 25-OCT-1999; | 99US-0161406. |
| PR | 26-OCT-1999; | 99US-0161359. |
| PR | 26-OCT-1999; | 99US-0161360. |
| PR | 26-OCT-1999; | 99US-0161361. |
| PR | 28-OCT-1999; | 99US-0161920. |
| PR | 28-OCT-1999; | 99US-0161992. |
| PR | 28-OCT-1999; | 99US-0161993. |
| PR | 29-OCT-1999; | 99US-0162142. |
| Query Match | | 69.1%; |
| Best Local Similarity | | 100.0%; |
| | | Score 1417; DB 21; Length 292; |
| | | Pred. No. 2.6e-131; |

| | | Matches | 270; | Conservative | 0; | Mismatches | 0; | Indels | 0; | Gaps | 0; |
|-----------|--|------------|----------|--------------|-----------|------------|-----------|--------|------|------|-------|
| Qy | 119 | MDQFHVSAFL | LEKGYQEA | BEITRRMGAG | MAKFCQ | EVETVDDY | DEYCHYVAG | LVL | 178 | | |
| Db | 1 | MDQFHVSAFL | LEKGYQEA | BEITRRMGAG | MAKFCQ | EVETVDDY | DEYCHYVAG | LVL | 60 | | |
| Qy | 179 | GLSKFLAAG | SEVLTPD | WEAISMSGL | FLOKTNII | RDYLEDINEI | PKSRMFWP | REI | WGKY | 238 | |
| Db | 61 | GLSKFLAAG | SEVLTPD | WEAISMSGL | FLOKTNII | RDYLEDINEI | PKSRMFWP | REI | WGKY | 120 | |
| Qy | 239 | ADKLEDKY | ENTNKSVQ | CLNEWVTN | ALMHIEDCL | KYMSLRDPS | IFR | FC | AI | QI | MAIGT |
| Db | 121 | ADKLEDKY | ENTNKSVQ | CLNEWVTN | ALMHIEDCL | KYMSLRDPS | IFR | FC | AI | QI | MAIGT |
| Qy | 299 | LALCYNNEQ | VFRGVK | LRGLTAKV | IDRTKT | MADVYGA | FDSCMLKTK | YDKNDP | N | NA | SKT |
| Db | 181 | LALCYNNEQ | VFRGVK | LRGLTAKV | IDRTKT | MADVYGA | FDSCMLKTK | YDKNDP | N | NA | SKT |
| Qy | 359 | LNRLEAVQ | KLCRDAGV | LQNRKSV | YNDK | GQP | 388 | | | | |
| Db | 241 | LNRLEAVQ | KLCRDAGV | LQNRKSV | YNDK | GQP | 270 | | | | |
| RESULT 10 | | | | | | | | | | | |
| AAG27739 | | | | | | | | | | | |
| ID | AAG27739 standard; Protein; 387 AA. | | | | | | | | | | |
| XX | AC AAG27739; | | | | | | | | | | |
| XX | DT 17-OCT-2000 (first entry) | | | | | | | | | | |
| DE | Arabidopsis thaliana protein fragment SEQ ID NO: 32695. | | | | | | | | | | |
| XX | KW Protein identification; signal transduction pathway; metabolic pathway; | | | | | | | | | | |
| KW | hybridisation assay; generic mapping; gene expression control; promoter; | | | | | | | | | | |
| KW | termination sequence. | | | | | | | | | | |
| XX | OS Arabidopsis thaliana. | | | | | | | | | | |
| XX | EP1033405-A2. | | | | | | | | | | |
| XX | PD 06-SEP-2000. | | | | | | | | | | |
| XX | PF 25-FEB-2000; 2000EP-0301439. | | | | | | | | | | |
| XX | PR 25-FEB-1999; 99US-0121825. | | | | | | | | | | |
| PR | 05-MAR-1999; 99US-0123180. | | | | | | | | | | |
| PR | 03-MAR-1999; 99US-0123548. | | | | | | | | | | |
| PR | 23-MAR-1999; 99US-0125788. | | | | | | | | | | |
| PR | 29-MAR-1999; 99US-0126264. | | | | | | | | | | |
| PR | 01-MAR-1999; 99US-0126785. | | | | | | | | | | |
| PR | 01-APR-1999; 99US-0127462. | | | | | | | | | | |
| PR | 06-APR-1999; 99US-0128234. | | | | | | | | | | |
| PR | 08-APR-1999; 99US-0128714. | | | | | | | | | | |
| PR | 16-APR-1999; 99US-0129845. | | | | | | | | | | |
| PR | 19-APR-1999; 99US-0130077. | | | | | | | | | | |
| PR | 21-APR-1999; 99US-0130449. | | | | | | | | | | |
| PR | 23-APR-1999; 99US-0130510. | | | | | | | | | | |
| PR | 23-APR-1999; 99US-0130891. | | | | | | | | | | |
| PR | 28-APR-1999; 99US-0131449. | | | | | | | | | | |
| PR | 30-APR-1999; 99US-0132048. | | | | | | | | | | |
| PR | 30-APR-1999; 99US-0132407. | | | | | | | | | | |
| PR | 04-MAY-1999; 99US-0132484. | | | | | | | | | | |
| PR | 05-MAY-1999; 99US-0132485. | | | | | | | | | | |
| PR | 06-MAY-1999; 99US-0132486. | | | | | | | | | | |
| PR | 06-MAY-1999; 99US-0132487. | | | | | | | | | | |
| PR | 07-MAY-1999; 99US-0132863. | | | | | | | | | | |
| PR | 11-MAY-1999; 99US-0134256. | | | | | | | | | | |
| PR | 14-MAY-1999; 99US-0134218. | | | | | | | | | | |
| PR | 14-MAY-1999; 99US-0134219. | | | | | | | | | | |
| PR | 14-MAY-1999; 99US-0134221. | | | | | | | | | | |
| PR | 14-MAY-1999; 99US-0134370. | | | | | | | | | | |
| PR | 18-MAY-1999; 99US-0134768. | | | | | | | | | | |

| | | |
|----|--------------|---------------|
| PR | 19-MAY-1999; | 99US-0134941. |
| PR | 20-MAY-1999; | 99US-0135124. |
| PR | 21-MAY-1999; | 99US-0135353. |
| PR | 24-MAY-1999; | 99US-0135629. |
| PR | 25-MAY-1999; | 99US-0136021. |
| PR | 27-MAY-1999; | 99US-0136392. |
| PR | 28-MAY-1999; | 99US-0136782. |
| PR | 01-JUN-1999; | 99US-0137222. |
| PR | 03-JUN-1999; | 99US-0137528. |
| PR | 04-JUN-1999; | 99US-0137502. |
| PR | 07-JUN-1999; | 99US-0137724. |
| PR | 08-JUN-1999; | 99US-0138094. |
| PR | 10-JUN-1999; | 99US-0138540. |
| PR | 14-JUN-1999; | 99US-0138847. |
| PR | 16-JUN-1999; | 99US-0139452. |
| PR | 16-JUN-1999; | 99US-0139453. |
| PR | 17-JUN-1999; | 99US-0139492. |
| PR | 18-JUN-1999; | 99US-0139455. |
| PR | 18-JUN-1999; | 99US-0139456. |
| PR | 18-JUN-1999; | 99US-0139457. |
| PR | 18-JUN-1999; | 99US-0139458. |
| PR | 18-JUN-1999; | 99US-0139459. |
| PR | 18-JUN-1999; | 99US-0139460. |
| PR | 18-JUN-1999; | 99US-0139461. |
| PR | 18-JUN-1999; | 99US-0139462. |
| PR | 18-JUN-1999; | 99US-0139463. |
| PR | 18-JUN-1999; | 99US-0139750. |
| PR | 18-JUN-1999; | 99US-0139763. |
| PR | 21-JUN-1999; | 99US-0139817. |
| PR | 22-JUN-1999; | 99US-0139899. |
| PR | 23-JUN-1999; | 99US-0140353. |
| PR | 23-JUN-1999; | 99US-0140354. |
| PR | 24-JUN-1999; | 99US-0140895. |
| PR | 28-JUN-1999; | 99US-0140823. |
| PR | 29-JUN-1999; | 99US-0140991. |
| PR | 30-JUN-1999; | 99US-0141287. |
| PR | 01-JUL-1999; | 99US-0141842. |
| PR | 02-JUL-1999; | 99US-0142154. |
| PR | 02-JUL-1999; | 99US-0142055. |
| PR | 06-JUL-1999; | 99US-0142390. |
| PR | 08-JUL-1999; | 99US-0142803. |
| PR | 09-JUL-1999; | 99US-0142920. |
| PR | 12-JUL-1999; | 99US-0142977. |
| PR | 13-JUL-1999; | 99US-0143542. |
| PR | 14-JUL-1999; | 99US-0143624. |
| PR | 15-JUL-1999; | 99US-0144005. |
| PR | 16-JUL-1999; | 99US-0144085. |
| PR | 16-JUL-1999; | 99US-0144086. |
| PR | 19-JUL-1999; | 99US-0144325. |
| PR | 19-JUL-1999; | 99US-0144331. |
| PR | 19-JUL-1999; | 99US-0144332. |
| PR | 19-JUL-1999; | 99US-0144333. |
| PR | 19-JUL-1999; | 99US-0144334. |
| PR | 20-JUL-1999; | 99US-0144352. |
| PR | 20-JUL-1999; | 99US-0144884. |
| PR | 21-JUL-1999; | 99US-0144814. |
| PR | 21-JUL-1999; | 99US-0145086. |
| PR | 21-JUL-1999; | 99US-0145088. |
| PR | 22-JUL-1999; | 99US-0145087. |
| PR | 22-JUL-1999; | 99US-0145089. |
| PR | 22-JUL-1999; | 99US-0145192. |
| PR | 23-JUL-1999; | 99US-0145218. |
| PR | 23-JUL-1999; | 99US-0145224. |
| PR | 26-JUL-1999; | 99US-0145276. |
| PR | 27-JUL-1999; | 99US-0145513. |
| PR | 27-JUL-1999; | 99US-0145518. |
| PR | 27-JUL-1999; | 99US-0145519. |

PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 69.1%; Score 1417; DB 21; Length 387;
Best Local Similarity 100.0%; Pred. No. 3.9e-131;
Matches 270; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|-----|---|-----|
| QY | 119 | MDQFHHVSAAFLELEKGYOEAEIEETRRMGAGMAKFICQEIVETVDYDDEYCHYVAGLVGL | 178 |
| DB | 1 | MDQFHHVSAAFLELEKGYOEAEIEETRRMGAGMAKFICQEIVETVDYDDEYCHYVAGLVGL | 60 |
| QY | 179 | GLSKLFLAAGSEVLTPDWEAISNSMGLFLOKTNIIRDYLEIDINEIPKSRMFEPREIWKY | 238 |
| DB | 61 | GLSKLFLAAGSEVLTPDWEAISNSMGLFLOKTNIIRDYLEIDINEIPKSRMFEPREIWKY | 120 |
| QY | 239 | ADKLEDLYEENTNKSVOCLNEMVTNALMHIEDCLKYWVSRLRDPISIFRFCAIQIQAIGT | 298 |
| DB | 121 | ADKLEDLYEENTNKSVOCLNEMVTNALMHIEDCLKYWVSRLRDPISIFRFCAIQIQAIGT | 180 |
| QY | 299 | LALCYNNEOVFGVVKLREGTLAKVIDRTKWADVYGAFYDFSCMLTKVKDKNDPNASKT | 358 |
| DB | 181 | LALCYNNEOVFGVVKLREGTLAKVIDRTKWADVYGAFYDFSCMLTKVKDKNDPNASKT | 240 |
| QY | 359 | LNRLEAVOKLCRDAGVLQNKRYSVNDKGOP | 388 |
| DB | 241 | LNRLEAVOKLCRDAGVLQNKRYSVNDKGOP | 270 |

RESULT 11
AAG37630
ID AAG37630 standard; Protein; 264 AA.
AC AAG37630;
XX XX
DT 18-OCT-2000 (first entry)
XX XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 46302.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX XX
PN EP1033405-A2.
PD 06-SEP-2000.
XX XX
PF 25-FEB-2000; 2000EP-0301439.
XX XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 99US-0132485.

| PR | 21-OCT-1999; | 99US-0160814. | PR | 19-APR-1999; | 99US-0130077. |
|----|--------------|---------------|----|--------------|---------------|
| PR | 21-OCT-1999; | 99US-0160815. | PR | 21-APR-1999; | 99US-0130449. |
| PR | 22-OCT-1999; | 99US-0160980. | PR | 23-APR-1999; | 99US-0130510. |
| PR | 22-OCT-1999; | 99US-0160981. | PR | 23-APR-1999; | 99US-0130891. |
| PR | 22-OCT-1999; | 99US-0160989. | PR | 28-APR-1999; | 99US-0131449. |
| PR | 25-OCT-1999; | 99US-0161404. | PR | 30-APR-1999; | 99US-0132048. |
| PR | 25-OCT-1999; | 99US-0161405. | PR | 30-APR-1999; | 99US-0132407. |
| PR | 25-OCT-1999; | 99US-0161406. | PR | 04-MAY-1999; | 99US-0132484. |
| PR | 26-OCT-1999; | 99US-0161359. | PR | 05-MAY-1999; | 99US-0132485. |
| PR | 26-OCT-1999; | 99US-0161360. | PR | 06-MAY-1999; | 99US-0132486. |
| PR | 26-OCT-1999; | 99US-0161361. | PR | 07-MAY-1999; | 99US-0132487. |
| PR | 28-OCT-1999; | 99US-0161320. | PR | 07-MAY-1999; | 99US-0132863. |
| PR | 28-OCT-1999; | 99US-0161392. | PR | 11-MAY-1999; | 99US-0134256. |
| PR | 28-OCT-1999; | 99US-0161393. | PR | 14-MAY-1999; | 99US-0134218. |
| PR | 28-OCT-1999; | 99US-0162142. | PR | 14-MAY-1999; | 99US-0134219. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 14-MAY-1999; | 99US-0134221. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 14-MAY-1999; | 99US-0134370. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 18-MAY-1999; | 99US-0134768. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 19-MAY-1999; | 99US-0134941. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 20-MAY-1999; | 99US-0135124. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 21-MAY-1999; | 99US-0135353. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 24-MAY-1999; | 99US-0135629. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 25-MAY-1999; | 99US-0136021. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 27-MAY-1999; | 99US-0136392. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 28-MAY-1999; | 99US-0136782. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 01-JUN-1999; | 99US-0137528. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 03-JUN-1999; | 99US-0137502. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 04-JUN-1999; | 99US-0137724. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 07-JUN-1999; | 99US-0138094. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 08-JUN-1999; | 99US-0138540. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 10-JUN-1999; | 99US-0138847. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 14-JUN-1999; | 99US-0139119. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 16-JUN-1999; | 99US-0139452. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 16-JUN-1999; | 99US-0139453. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 17-JUN-1999; | 99US-0139492. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 18-JUN-1999; | 99US-0139454. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 18-JUN-1999; | 99US-0139455. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 18-JUN-1999; | 99US-0139456. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 18-JUN-1999; | 99US-0139457. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 18-JUN-1999; | 99US-0139458. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 18-JUN-1999; | 99US-0139459. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 18-JUN-1999; | 99US-0139460. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 18-JUN-1999; | 99US-0139461. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 18-JUN-1999; | 99US-0139462. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 18-JUN-1999; | 99US-0139463. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 18-JUN-1999; | 99US-0139464. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 18-JUN-1999; | 99US-0139465. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 18-JUN-1999; | 99US-0139466. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 18-JUN-1999; | 99US-0139467. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 18-JUN-1999; | 99US-0139468. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 18-JUN-1999; | 99US-0139469. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 18-JUN-1999; | 99US-0139470. |
| PR | 29-OCT-1 | | | | |

```
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145132.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145921.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147132.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147433.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.

PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 62.1%; Score 1274; DB 21; Length 359;
Best Local Similarity 100.0%; Pred. No. 4.7e-117;
Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 147 MGAGMAKFCQEVETVDDYDEYHYVAGLVGLGSLKFLAAGSEVLTPDWEAISNSGLF 206
Db 1 MGAGMAKFCQEVETVDDYDEYHYVAGLVGLGSLKFLAAGSEVLTPDWEAISNSGLF 60

Qy 207 LQKTNIIIRDYLEINEIPKSRMFWPREIWGKYADKLEDKYEENTNKSVOCLNEMVTNAL 266
Db 61 LQKTNIIIRDYLEINEIPKSRMFWPREIWGKYADKLEDKYEENTNKSVOCLNEMVTNAL 120

Qy 267 MHIEDCLKYMWSLRDPISIFRFCAIPQIMAITGLALCYNNEQVFRGVVKKLRGLTAKVIDR 326
Db 121 MHIEDCLKYMWSLRDPISIFRFCAIPQIMAITGLALCYNNEQVFRGVVKKLRGLTAKVIDR 180

Qy 327 TKTMDVYGAFYDFSCMLKTKVDKNDPNASKTUNRLEAVOKLORDAGVLQNRKSYVNDKG 386
Db 181 TKTMDVYGAFYDFSCMLKTKVDKNDPNASKTUNRLEAVOKLORDAGVLQNRKSYVNDKG 240

Qy 387 QP 388
Db 241 QP 242

RESULT 13
AAG37631
ID AAG37631 standard; Protein; 260 AA.
XX
AC AAG37631;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 46303.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
```

us-10-024-130a-6.rag

Thu Nov 13 10:01:34 2003

PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135253.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 21-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140931.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149923.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.

```

PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159291.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160774.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 61.1%; Score 1253; DB 21; Length 260;
Best Local Similarity 100.0%; Pred. No. 3.5e-115;
Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 151 MAKFIQCEVETVDDYDEYCHYVAGLVGLGSLKFLAAGSEVLTDPWEAINSMGLFLQKT 210
Db 1 MAKFIQCEVETVDDYDEYCHYVAGLVGLGSLKFLAAGSEVLTDPWEAINSMGLFLQKT 60

Qy 211 NIIRDYLEDINEIPKSRMFPRFETWGWYADKLEDLKYEENTNKSVOCLNEMVTNALMHIE 270
Db 61 NIIRDYLEDINEIPKSRMFPRFETWGWYADKLEDLKYEENTNKSVOCLNEMVTNALMHIE 120

Qy 271 DCLKYMVSLRDPSTFRFCAIPQIMAGTLCALCYNNEQVFRGVVWKLRRGLTAKVIDRRTKM 330
Db 121 DCLKYMVSLRDPSTFRFCAIPQIMAGTLCALCYNNEQVFRGVVWKLRRGLTAKVIDRRTKM 180

Qy 331 ADVYGAFYDFSCMLTKTKVDKNDPNASKTLNRLNLEAVQKLCRDAGVLQNRKSYVNDKGOP 388
Db 181 ADVYGAFYDFSCMLTKTKVDKNDPNASKTLNRLNLEAVQKLCRDAGVLQNRKSYVNDKGOP 238

RESULT 14
ABBS7061
ID ABB57061 standard; Protein; 416 AA.
XX AC ABB57061;
XX XX
XX 07-MAR-2002 (first entry)
XX DE
XX Mouse ischaemic condition related protein sequence SEQ ID NO:118.
XX KW Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
XX KW vasospastic ischaemia; ischaemic condition; ischaemic disease.
XX OS Mus musculus.
XX XX
XX WO200188188-A2.
XX PN
XX 22-NOV-2001.
XX PD

```

```

XX 18-MAY-2001; 2001WO-JP04192.
XX 18-MAY-2000; 2000JP-0145977.
XX (UYN1-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
XX WPI; 2002-034733/04.
XX N-PSDB; ABI99277.
XX Examining the ischemic condition (e.g. occlusive ischemia) by measuring
XX expression levels of particular genes defined in the specification or
XX by determining the expression profile of a gene group comprising these
XX genes -
XX Claim 2; Page 345-346; 2690pp; English.
XX The present invention describes a method for examining ischaemic
XX conditions, comprising measuring the expression levels of particular
XX genes (I) in a test sample or determining the expression profile of a
XX gene group in the sample comprising genes selected from (I). The method
XX is useful for examining the ischaemic condition (e.g. compressive
XX ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
XX the protein sequences in ABB57020 to ABB57374) or by determining the
XX expression profile of a gene group comprising these genes. The
XX expression levels or expression profiles produced by these genes are
XX used as an indicator when screening for ischaemic condition-improving
XX drugs or therapeutics for ischaemic diseases. ABB57374 and ABB57374
XX represent PCR primers for a mouse ischaemic condition related sequence,
XX which are used in the exemplification of the present invention.
XX SQ Sequence 416 AA;

Query Match 41.3%; Score 848; DB 23; Length 416;
Best Local Similarity 46.9%; Pred. No. 7.2e-75;
Matches 169; Conservative 70; Mismatches 108; Indels 12; Gaps 5;

Qy 8 LRYDDIIPLLKKRA-----IEKAEKQIPPEPHWGFYSMLHKVSRFSLSVLIQQLNTEL 62
Db 7 LGHPEEFYNLLRFRMGRRNFIPKMD-QDSLSSSLKTCYKYLNQTSRFAAVIQALDGD 65

Qy 63 RNACVCFYLVRLALDTVEDDTSIPTDEKVPILIAFHRHIYDTDWHYSCGTKEYKILMDQF 122
Db 66 RHAICVYLVRLALDTVEDDMSISVEKKIPLCNFHTFLYDPWRFTESKEKORQVLEDP 125

Qy 123 HHVSAAPLELEKGYQEAIEIEITRMGAGMAKFTICQEVETVDDYDEYCHYVAGLVGLSK 182
Db 126 PTISLEFRNLAEKYQTVIDDICHQMGCGMAEFVDKDVTSKQMDKYCHYVAGLVGIGLSR 185

Qy 183 LFLAASE--VLTDPWEAINSMGLFLOKTNIRDYLEDINEIPKSRMFPRFETWGWYAD 240
Db 186 LFSASEFEDPIVGEDIEC-ANSNGLFLOKTNIRDYLEDQEE---GRKFPQEWGRIYK 241

Qy 241 KLEDLKYEENTNKSVOCLNEMVTNALMHIEDCLKYMVSLRDPSTFRFCAIPQIMAGTTLA 300
Db 242 KLEDFAPKPNVDVAVOCLNELITNTLOHI PDVLTYSRLNQSVNFCAIPQVMAIATLA 301

Qy 301 LCYNNEQVFRGVVWKLRRGLTAKVIDRRTKWADYVGFYDFSCMLTKTKVDKNDPNASKT 358
Db 302 ACYNNQVFKGVVKIRKQAVTLMMDATNMPAVKAIYQVIEEYHRIIPNSDPSSSKT 359

RESULT 15
AARS2606
ID AARS2606 standard; Protein; 417 AA.
XX AC AARS2606;
XX XX
XX 05-DEC-1994 (first entry)
XX DT
XX XX

```

Thu Nov 13 10:01:34 2003

DE Human squalene synthase.
XX Squalene synthase; sterol; metabolism; hypercholesterolemia;
KW atherosclerosis; treatment; therapy; prevention.
XX Homo sapiens.
XX GB2272442-A.
XX 18-MAY-1994.
XX 09-NOV-1993; 93GB-0023035.
XX 11-NOV-1992; 92GB-0023610.
XX (ZENE) ZENECA LTD.
XX Charles AD;
XX WPI; 1994-146577/18.
DR N-PSDB; AAQ62598.
XX New recombinant human squalene synthase - used for screening for
PT inhibitors which can be used in the treatment or prevention of
PT high cholesterol levels
XX Claim 3; Page 36-38; 53pp; English.
XX Recombinantly produced human squalene synthase may be used as a
CC source of enzyme in a non-sterol metabolising host for enzymatic
CC studies or for the screening of compounds to identify inhibitors.
CC Selective inhibition of human squalene synthase can be used to lower
CC intracellular cholesterol levels and provide improved treatment for,
CC and prevent, hypercholesterolemia and atherosclerosis. Human
CC squalene synthase can also be used for the production of antibodies.
XX Sequence 417 AA;
SQ
Query Match 40.0%; Score 820; DB 15; Length 417;
Best Local Similarity 45.2%; Pred. No. 4.2e-72; Indels 10; Gaps 4;
Matches 165; Conservative 69; Mismatches 121;
QY 8 LRYPDIDYPLKWM---KRAIEKAEKQIPPEPHWGFCSYMLHKVSRSPSLVIOQLNTELR 63
DB 7 LGHPEFYNLVRIGGKRVKPKMDQDSSLSSLKTCYKYLNOTSRSPAFAVIALDGEWR 66
QY 64 NAVCVFVLRLALDTVEDDTSIPTDEKVPILIAFHRHIYDTDWHYSCGTKEYKILMDQPH 123
DB 67 NAVCIFVLRLALDTLEDDMTISVEKKVPLLNHFNHFLYQPDWRFMESKEKDRQVLEDFP 126
QY 124 HVSAARLELEKGYQEAIEEITRRMGAGMAKFIQCEVETVDDYDEYCHYVAGLVGLSL 183
DB 127 TISLEFRNLAEKYQTVIADICRMGIGMAEFLDKHVTSEQENDKYCHYVAGLVGLSRL 186
QY 184 FLAAGSE--VLTDPWEAISMGFLQKTNIIIRDYLEINEIPKSRMFWPREIWGYADK 241
DB 187 FSASEFEDPLVGEDTER-ANSMGLFLQKTNIIIRDYLEDDQ--GGREFWPQEVMSRYVK 242
QY 242 LEDLKYENTNKSVOCLNEMVTNALMHIEDCLKVMYSLRDPISFRCAIPQIMAGTLAL 301
DB 243 LGDFAKPENIDLAQQCLNELITNALHHIPDVITYLSLRNQSVFNCAIPQWAIATLAA 302
QY 302 CYNNEQVFRGVVKLRRLGTAKVIDRTKTMADVYGAFYDFSCMLKTKVDKNDPNASKTLNR 361
DB 303 CYNNOQVFKGAVIRKQGAVTLMMDATNMPAVKAIYYQYMEIIVHRIPDSDPSSSKTRQI 362
QY 362 LEAVQ 366
DB 363 ISTIR 367

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 12, 2003, 16:28:06 ; Search time 26 Seconds
(without alignments)
631.408 Million cell updates

```

PERFECT SCORE: 2031
Sequence: 1 MGS LGT M L R Y P D D I Y P L L K M ..... CRDAGV LONR KSYVNDKGOP 388

```

Scoring table: BLOSUM62

Learning curve: $\text{Gapop } 10.0$, $\text{Gapext } 0.5$

Searched: 328717 seqs. 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000
Maximum DB seq length: 200000000

.....

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

```
Database : Issued Patents AA:
1. /cgn2 6/ptofata/11aa/5A COMB res.*
```

```

1: /cgn2_6/ptodaca/1/taa/5A_COMB.per:
2: /cgn2_6/ptodaca/1/taa/5B_COMB.per:

```

```
2: /cqn2_6/ptodata/1/1aa/35_COMB.per:
3: /cqn2_6/ptodata/1/1aa/6A_COMB.per:*
```

4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*

5: /cgn2_6/ptodata/1/iaa/pCTUS_COMB.pe

6: /cgn2_6/ptodata/1/iaa/backfiles1.pe

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | ID | | | Description |
|------------|-------|-------------|--------|----|--------------------|--------------------|
| | | | Length | DB | ID | |
| 1 | 1618 | 78.9 | 411 | 1 | US-08-310-693-2 | Sequence 2, Appli |
| 2 | 1618 | 78.9 | 411 | 5 | PTC-US955-11280-2 | Sequence 2, Appli |
| 3 | 820 | 40.0 | 417 | 1 | US-08-351-981-6 | Sequence 6, Appli |
| 4 | 814 | 39.7 | 417 | 1 | US-08-351-981-7 | Sequence 7, Appli |
| 5 | 722 | 35.2 | 444 | 1 | US-08-351-981-2 | Sequence 2, Appli |
| 6 | 715 | 34.9 | 444 | 1 | US-08-351-981-8 | Sequence 8, Appli |
| 7 | 709.5 | 34.6 | 441 | 4 | US-09-625-188-4 | Sequence 4, Appli |
| 8 | 700.5 | 34.2 | 460 | 1 | US-08-351-981-4 | Sequence 4, Appli |
| 9 | 695.5 | 33.9 | 460 | 1 | US-08-351-981-9 | Sequence 9, Appli |
| 10 | 115 | 5.6 | 296 | 1 | US-07-783-705A-5 | Sequence 5, Appli |
| 11 | 108 | 5.3 | 410 | 1 | US-08-579-667-2 | Sequence 2, Appli |
| 12 | 107.5 | 5.2 | 309 | 1 | US-08-331-004A-2 | Sequence 2, Appli |
| 13 | 107.5 | 5.2 | 309 | 5 | PTC-US955-13937A-2 | Sequence 2, Appli |
| 14 | 104.5 | 5.1 | 308 | 1 | US-08-095-726-6 | Sequence 6, Appli |
| 15 | 104.5 | 5.1 | 308 | 1 | US-08-096-043-6 | Sequence 6, Appli |
| 16 | 104.5 | 5.1 | 308 | 1 | US-08-093-577-6 | Sequence 6, Appli |
| 17 | 104.5 | 5.1 | 308 | 1 | US-08-098-623A-6 | Sequence 6, Appli |
| 18 | 104 | 5.1 | 410 | 1 | US-08-579-667-8 | Sequence 8, Appli |
| 19 | 103 | 5.0 | 410 | 1 | US-08-579-667-6 | Sequence 6, Appli |
| 20 | 102 | 5.0 | 892 | 4 | US-08-579-667-6 | Sequence 6, Appli |
| 21 | 98.5 | 4.8 | 413 | 1 | US-08-579-667-4 | Sequence 8164, Ap |
| 22 | 97 | 4.7 | 733 | 4 | US-09-198-452A-791 | Sequence 4, Appli |
| 23 | 96.5 | 4.7 | 402 | 3 | US-09-180-342-3 | Sequence 791, App |
| 24 | 96.5 | 4.7 | 580 | 3 | US-09-234-393-48 | Sequence 3, Appli |
| 25 | 96.5 | 4.7 | 580 | 3 | US-09-865-171-48 | Sequence 48, Appli |
| 26 | 96.5 | 4.7 | 581 | 3 | US-09-234-393-20 | Sequence 48, Appli |
| 27 | 96.5 | 4.7 | 581 | 3 | US-09-234-393-44 | Sequence 20, Appli |

QY 61 ELRNAVCVYLVRALDVTEDDTSIPTDEKVPILIAFHRHIYDTHWYSCGTKEYKILMD 120
DB 61 ELRDVAVCI FYLVRALDVTEDDTSIPTDVKVPILISPHQHVYDREWHFSCGTKEYKVLMD 120
QY 121 QFHVSAFALEKGYQEAIBEITRRMGAGMAKFCOEVEVTDYDDEYCHYVAGLVGL 180
DB 121 QFHVSTAFLELRKHYYQQAIEDITRRMGAGMAKFCOEVEVTDYDDEYCHYVAGLVGL 180
QY 181 SKFLAAGSEVLTDPWEAISNSMGLFLOKTNIIIRDYLEINEIPKSRMFWPREIWGYAD 240
DB 181 SKLPHASEKEDLASD--SLNSMGLFLOKTNIIIRDYLEINEIPKSRMFWPREIWSKYVN 238
QY 241 KLEDKYBENTNKSVOCLNEMVTNALMHIEDCLKYVUSLRDPSIFRCAIPQVMAIGTLA 300
DB 239 KLEELKYEDNSAKAVQCLNDVMTNALSHVEDCLTYMSALRDPISIFRCAIPQVMAIGTLA 298
QY 301 LCYNNEQVFRGVVKKRRGLTAKVIDRTKTMADVYGAFYDFSCMLKTKVDKNDPNASKTLN 360
DB 299 MCYDNIEVFRGVVKKRRGLTAKVIDRTKTMADVYGAFYDFSCMLKSKVNNNDPNATKTLK 358
QY 361 RLEAVQKLCRDAGVLQNRKSYV 382
DB 359 RLEVILKTCRDSGTLNKRKSYI 380

RESULT 2
PCT-US95-11280-2
; Sequence 2, Application PC/TUS9511280
; GENERAL INFORMATION:
; APPLICANT: Hanley, Kathleen M.
; APPLICANT: Nicolas, Olivier
; APPLICANT: Hellmann, Gary M.
; TITLE OF INVENTION: DNA Sequence encoding squalene
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: R. J. Reynolds Tobacco Company
; STREET: 950 Reynolds Blvd., P. O. Box 1487
; CITY: Winston-Salem
; STATE: North Carolina
; COUNTRY: U. S. A.
; ZIP: 27102-1487
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/11280
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Borschke, August J.
; REGISTRATION NUMBER: 30,539
; REFERENCE/DOCKET NUMBER: cc-212A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 910-741-5491
; TELEFAX: 910-741-5449
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-11280-2

Query Match 78.9%; Score 1618; DB 5; Length 411;
Best Local Similarity 77.5%; Pred. No. 5.3e-158;
Matches 296; Conservative 45; Mismatches 39; Indels 2; Gaps 1;
QY 1 MGSIGTMLRYPDDLYPLVKLAAARHAEKQIIPPENWGFQCYSMHLKYSRSPALVIOQLPV 60

DB 1 MGSRLAILKNPEDLYPLVKLAAARHAEKQIIPPENWGFQCYSMHLKYSRSPALVIOQLPV 60
QY 61 ELRNAVCVYLVRALDVTEDDTSIPTDEKVPILIAFHRHIYDTHWYSCGTKEYKILMD 120
DB 61 ELRDVAVCI FYLVRALDVTEDDTSIPTDVKVPILISPHQHVYDREWHFSCGTKEYKVLMD 120
QY 121 QFHVSAFALEKGYQEAIBEITRRMGAGMAKFCOEVEVTDYDDEYCHYVAGLVGL 180
DB 121 QFHVSTAFLELRKHYYQQAIEDITRRMGAGMAKFCOEVEVTDYDDEYCHYVAGLVGL 180
QY 181 SKFLAAGSEVLTDPWEAISNSMGLFLOKTNIIIRDYLEINEIPKSRMFWPREIWGYAD 240
DB 181 SKLPHASEKEDLASD--SLNSMGLFLOKTNIIIRDYLEINEIPKSRMFWPREIWSKYVN 238
QY 241 KLEDKYBENTNKSVOCLNEMVTNALMHIEDCLKYVUSLRDPSIFRCAIPQVMAIGTLA 300
DB 239 KLEELKYEDNSAKAVQCLNDVMTNALSHVEDCLTYMSALRDPISIFRCAIPQVMAIGTLA 298
QY 301 LCYNNEQVFRGVVKKRRGLTAKVIDRTKTMADVYGAFYDFSCMLKTKVDKNDPNASKTLN 360
DB 299 MCYDNIEVFRGVVKKRRGLTAKVIDRTKTMADVYGAFYDFSCMLKSKVNNNDPNATKTLK 358
QY 361 RLEAVQKLCRDAGVLQNRKSYV 382
DB 359 RLEVILKTCRDSGTLNKRKSYI 380

RESULT 3
US-08-351-981-6
; Sequence 6, Application US/08351981
; Patent No. 5589372
; GENERAL INFORMATION:
; APPLICANT: Robinson, Gordon W.
; TITLE OF INVENTION: Squalene Synthetase
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burton Rodney
; STREET: P.O. Box 4000
; CITY: Princeton
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 08543-4000
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/351,981
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/911,835
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Gaul, Timothy J.
; REGISTRATION NUMBER: 33,111
; REFERENCE/DOCKET NUMBER: DC7a
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (609) 252-5901
; TELEFAX: (609) 252-4526
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 417 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-351-981-6

Query Match 40.0%; Score 820; DB 1; Length 417;
Best Local Similarity 45.2%; Pred. No. 7e-76;
Matches 165; Conservative 69; Mismatches 121; Indels 10; Gaps 4;

Db 178 YCHYVAGLVGDLTHLIMLAKSPGLYDSDPOLYESMGLFLQKTNIIIRDYASDLAD --- 234
Qy 226 SRMFWEIRMGKYADKLEDLYKENTNKSVOCLNEMVTNMAHIEDCLKYMVSLRPPSIF 285
Db 235 GRSEFMPKEIWSHWYADDLASFSPKENATAGVYVCHLVLNAGHQHVHVTYLASLRQSSP 294
Qy 286 RFCAIQMAIGTALCYNNEQVFRGVVKLRGLTAKVIDRTKTMADYCAFYDFSCMLK 345
Db 295 QFCAIQVMAIATLALVFGNERVLOTSKIRKGTTCVYLILKSRFTQGCVEIFEHYLRDIR 354
Qy 346 TKVDKNDPNASK 357
Db 355 KRLTVADPNYLK 366

RESULT 8
US-08-351-981-4
; Sequence 4, Application US/08351981
; Patent No. 5589372
; GENERAL INFORMATION:
; APPLICANT: Robinson, Gordon W.
; TITLE OF INVENTION: Squalene Synthetase
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burton Rodney
; STREET: P.O. Box 4000
; CITY: Princeton
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 08543-4000
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/351,981
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/911,835
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Gaul, Timothy J.
; REGISTRATION NUMBER: 33,111
; REFERENCE/DOCKET NUMBER: DC7a
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (609) 252-5901
; TELEFAX: (609) 252-4526
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 460 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-351-981-4

Query Match 34.2%; Score 700.5; DB 1; Length 460;
Best Local Similarity 44.5%; Pred. No. 1.6e-63;
Matches 147; Conservative 60; Mismatches 106; Indels 17; Gaps 5;
Qy 40 CYSMLHKVRSFSLVLIQQLNTELRNAVCFYLVLRALDVTVEDDTSIPTDEKVPILIAFHR 99
Db 38 CYQLLDWTSRSPAVVVKLPNGIRQAVMIFVYLVRLGLDVTVEDDWTLPDCKKLPILRDFYK 97
Qy 100 HIYDTHWHY--SCGTKEYKILMQDFHHVSAFAFLEKGYQEAIEEITRRMGAGNAKFIQ 157
Db 98 TIEVEGTFNESGPNKDRQLLVFQVVIKEYLNLSEGYRNVISNITKEMGDGMAYVASL 157
Qy 158 -----EVTVDYDDEYCHYVAGLVGLGSKLFLAAGSEVLTPDW---EAINSMGLFL 207
Db 158 AEKNDGFSVETIDFNKYCHYVAGLVGLGSRUF--AQSKLEDPDLAHSQAISNSLGLFL 215

Qy 208 QKTNIIROYLEDINEIPKSRMFWPREIMWKYADKLEDLYKENTNKSVOCLNEMVTNMA 267
Db 216 QKVNIIIRDYREDPDD---NRHFWPREIMWSKYTSSFGDLCPLPDNSEKALECLSDMTANALT 272
Qy 268 HTIEDCLKYVUSLRDPSIFRFCAIQMAIGTALCYNNEQVFRGVVKLRGLTAKVIDRT 327
Db 273 HATDALVYLSQLTKQIFNFCAIQVMAIATLAAPFRNDPVDFTQNVKIRKQAVQIILHS 332
Qy 328 KTMADYVYGFYDFSCMLKTKVDKNDPNASK 357
Db 333 VNLKNVCDLFLRYTRDIHYKNTKPDNFKL 362

RESULT 9
US-08-351-981-9
; Sequence 9, Application US/08351981
; Patent No. 5589372
; GENERAL INFORMATION:
; APPLICANT: Robinson, Gordon W.
; TITLE OF INVENTION: Squalene Synthetase
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burton Rodney
; STREET: P.O. Box 4000
; CITY: Princeton
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 08543-4000
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/351,981
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/911,835
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Gaul, Timothy J.
; REGISTRATION NUMBER: 33,111
; REFERENCE/DOCKET NUMBER: DC7a
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (609) 252-5901
; TELEFAX: (609) 252-4526
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 460 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-351-981-9

Query Match 33.9%; Score 695.5; DB 1; Length 460;
Best Local Similarity 44.8%; Pred. No. 5.3e-63;
Matches 148; Conservative 59; Mismatches 107; Indels 17; Gaps 5;
Qy 40 CYSMLHKVRSFSLVLIQQLNTELRNAVCFYLVLRALDVTVEDDTSIPTDEKVPILIAFHR 99
Db 38 CYQLLDWTSRSPAVVVKLPNGIRQAVMIFVYLVRLGLDVTVEDDWTLPDCKKLPILRDFYK 97
Qy 100 HIYDTHWHY--SCGTKEYKILMQDFHHVSAFAFLEKGYQEAIEEITRRMGAGNAKFIQ 157
Db 98 TIEVEGTFNESGPNKDRQLLVFQVVIKEYLNLSEGYRNVISNITKEMGDGMAYVASL 157
Qy 158 -----EVTVDYDDEYCHYVAGLVGLGSKLFLAAGSEVLTPDW---EAINSMGLFL 207
Db 158 AEKNDGFSVETIDFNKYCHYVAGLVGLGSRUF--AQSKLEDPDLAHSQAISNSLGLFL 215
Qy 208 QKTNIIIRDYLEDINEIPKSRMFWPREIMWKYADKLEDLYKENTNKSVOCLNEMVTNMA 267

157 QEVETVDDYDEYCHVAGLVGLSLKFLAAGSEVLTDPDWEAISNSMGLFLOKNIIRDY 216
; Sequence 2, Application US/08579667
; Patent No. 5705624
; GENERAL INFORMATION:
; APPLICANT: Fitzmaurice, Wayne P.
; APPLICANT: Hellmann, Gary M.
; APPLICANT: Grill, Laurence K.
; APPLICANT: Kumagai, Monto H.
; APPLICANT: Della-Cioppa, Guy R.
; TITLE OF INVENTION: DNA SEQUENCES ENCODING ENZYMES USEFUL IN
; TITLE OF INVENTION: PHYTOENE BIOSYNTHESIS
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Virginia C. Bennett
; STREET: 1211 East Morehead Street, PO Drawer 34009
; CITY: Charlotte
; STATE: No. 5705624th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/579,667
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Bennett, Virginia C.
; REGISTRATION NUMBER: 37,092
; REFERENCE/DOCKET NUMBER: 627-196
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-420-2200
; TELEFAX: 919-881-3175
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 410 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-579-667-2

Query Match 5.3%; Score 108; DB 1; Length 410;
Best Local Similarity 21.3%; Pred. No. 0.013;
Matches 58; Conservative 48; Mismatches 108; Indels 58; Gaps 11;

8 LRYDDIYPLLMKRAIEKAEKQIPPEPH-----WGFCYSMLHKVSRFSFLVIOQLNT 60
94 LRSTDDL-----EVKPEIPLQANLSLSEAYDRCSEVCAEYAKTYLGTMLMTP 142

61 ELRNACVFLVLRALDVTEDDTSIPTDEKVPILIAFHRIHYDTDMHYSQCTKEYKILMD 120
143 ERRRAIWAIVYWCRTBELVDGPNA-----SHITPQ-----GLDRWEDLLE 183

121 Q-----FHVSAFALEKGYQBAIEITRRMGAGMAKFCQ-EVETVDDYDEYCHV 173
184 DVFSGRPFMDLAALSDDTVSKFPVDIQPF-RDMIEGRMDLRKGRYRNFDELYCYVA 242

174 GLVGL-GLSKFLAAGSEVLTDPDWEAISNSMGLFLOKNIIRDYLEDINEIPKGRMFWPR 232
243 GTVGLMSVPIMGIAFDPSKATTENVYNAALGIANQLNLRDVGEDAR-----RGVYLPQ 299

216 QKVNIIIRDYREDFDD---NRHFWPREIWSKYTSFGDLCTPCNSEKALECLSDMTANALT 272
268 HIEDCLAYMSLRDPSIFRCAIPQINAIQTALALCYNNEQVFRGVVVKLRGLTAKVIDRT 327
273 HATDALVLSQKLTQEIFNCAIPQVNAIATLAVERNPLVFTQNVKIRKQVAVQIILHS 332

328 KTMADVYGAFYDFSCMLKTKVDKNDPNASK 357
333 VNLKNVCDLFURYTRDIHYKNTKPDNPLK 362

RESULT 10
US-07-783-705A-5
; Sequence 5, Application US/07783705A
; Patent No. 5429939
; GENERAL INFORMATION:
; APPLICANT: Misawa, No. 5429939ihiko
; APPLICANT: Kobayashi, Kazuo
; APPLICANT: Nakamura, Katsumi
; APPLICANT: Yamano, Shigeyuki
; TITLE OF INVENTION: DNA SEQUENCES USEFUL FOR THE
; TITLE OF INVENTION: SYNTHESIS OF CAROTENOIDS
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ladas & Parry
; STREET: 26 West 61 Street
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10023
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720Kb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: N/A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/783,705A
; FILING DATE: 19911023
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER: JP 1-103078
; FILING DATE: 21-APR-1989
; APPLICATION NUMBER: JP 2-53225
; FILING DATE: 05-MAR-1990
; APPLICATION NUMBER: US 07/519,011
; FILING DATE: 19-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Schwadron, Janet I.
; REGISTRATION NUMBER: 33,778
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-708-1935
; TELEFAX: 212-246-5959
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 296 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-783-705A-5

Query Match 5.6%; Score 115; DB 1; Length 296;
Best Local Similarity 21.8%; Pred. No. 0.0015;
Matches 51; Conservative 43; Mismatches 102; Indels 38; Gaps 9;

48 SRFSFLVIOQLNTELNRNACVFLVLRALDVTEDDTSIPTDEKVPILIAFHRIHYDTDM- 106
5 SKSFATASKLFDKATRTASVLMYAWCRHCDVDVDDDTLGLFGQARQALQTPGRLMQLMK 64

107 --HYSCCTKEYKILMDQFHVSA-----AFLELEKGYQBAIEITRRMGAGMAKFC 156
65 TRQYAGSQMHPEAFARQFVAMAHDIAPAYAFDHLF-GFAMDVRE----- 109

Qy 233 EWGK-----YADKLEDLKYENTNKSQV 256
 Db 300 DELAHAGLSDDIPAGKVTD-KWRSFMKKQIQ 330

RESULT 12

US-08-331-004A-2
 ; Sequence 2, Application US/08331004A

; Patent No. 5618988

; GENERAL INFORMATION:

; APPLICANT: Hauptmann, Randal

; APPLICANT: Eschenfeldt, William H

; APPLICANT: English, Jami

; APPLICANT: Brinkhaus, Friedhelm L

; TITLE OF INVENTION: Enhanced Carotenoid Accumulation

; TITLE OF INVENTION: in Storage Organs of Genetically

; TITLE OF INVENTION: Engineered Plants

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Amoco Corporation, Law Dept

; STREET: 55 Shuman Boulevard, Suite 600

; CITY: Naperville

; STATE: IL

; COUNTRY: USA

; ZIP: 60563-8437

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.24

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/331.004A

; FILING DATE:

; CLASSIFICATION: 800

; ATTORNEY/AGENT INFORMATION:

; NAME: Galloway, No. 5618988val B

; TELEPHONE: 7087172447

; TELEFAX: 7087172430

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 309 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-331-004A-2

Query Match 5.2%; Score 107.5; DB 1; Length 309;

Best Local Similarity 25.1%; Pred. No. 0.0092;

Matches 48; Conservative 27; Mismatches 69; Indels 47; Gaps 7;

Qy 48 SRSFSLVIQQLNTELRNACVFLVLRALDVTDEDD-----TSIPTDEKVPILIAFHRHI 101

Db 18 SKSFATAAKLFDPATRSVLMYTWCRHCDVDIDQTHGFASAAAEEATQRLARLRTL 77

Qy 102 YDTDMHYSCTGKEYKILMDQFHVSAAF--LELEKGYQEAIEE-----ITRRMGA---- 149

Db 78 -----TLAAFEAGEMQDPAAFAQFQVALTHGITPRMALDHL 114

Qy 150 GMAKFICQ-EVETVDDYDEYCHYVAGLVGLSKLFLAAGSEVLTDPWEAISMSGLFLQ 208

Db 115 GFAMDVAQTRYVTFTEDTLRYCYHVAGVGLMMARVMGVDRVLDL-----ACDLGLAFQ 169

Qy 209 KTNIIIRDYLE 219

Db 170 LTNIAIDIID 180

RESULT 13

PCT-US95-13937A-2

; Sequence 2, Application PC/TUS9513937A

; GENERAL INFORMATION:

; APPLICANT: Hauptmann, Randal

; APPLICANT: Eschenfeldt, William H
 ; APPLICANT: English, Jami
 ; APPLICANT: Brinkhaus, Friedhelm L
 ; TITLE OF INVENTION: Enhanced Carotenoid Accumulation
 ; TITLE OF INVENTION: in Storage Organs of Genetically
 ; TITLE OF INVENTION: Engineered Plants
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Amoco Corporation, Law Dept

; STREET: 55 Shuman Boulevard, Suite 600

; CITY: Naperville

; STATE: IL

; COUNTRY: USA

; ZIP: 60563-8437

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.24

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US95/13937A

; FILING DATE:

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Galloway, Norval B

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 7087172447

; TELEFAX: 7087172430

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 309 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

PCT-US95-13937A-2

Query Match 5.2%; Score 107.5; DB 5; Length 309;

Best Local Similarity 25.1%; Pred. No. 0.0092;

Matches 48; Conservative 27; Mismatches 69; Indels 47; Gaps 7;

Qy 48 SRSFSLVIQQLNTELRNACVFLVLRALDVTDEDD-----TSIPTDEKVPILIAFHRHI 101

Db 18 SKSFATAAKLFDPATRSVLMYTWCRHCDVDIDQTHGFASAAAEEATQRLARLRTL 77

Qy 102 YDTDMHYSCTGKEYKILMDQFHVSAAF--LELEKGYQEAIEE-----ITRRMGA---- 149

Db 78 -----TLAAFEAGEMQDPAAFAQFQVALTHGITPRMALDHL 114

Qy 150 GMAKFICQ-EVETVDDYDEYCHYVAGLVGLSKLFLAAGSEVLTDPWEAISMSGLFLQ 208

Db 115 GFAMDVAQTRYVTFTEDTLRYCYHVAGVGLMMARVMGVDRVLDL-----ACDLGLAFQ 169

Qy 209 KTNIIIRDYLE 219

Db 170 LTNIAIDIID 180

RESULT 14

US-08-095-726-6

; Sequence 6, Application US/08095726

; Patent No. 5530188

; GENERAL INFORMATION:

; APPLICANT: Ausich, Rodney L

; APPLICANT: Brinkhaus, Friedhelm L

; APPLICANT: Mukharji, Indrani

; APPLICANT: Proffitt, John H

; APPLICANT: Yarger, James G

; APPLICANT: Yen, Hwei-Che B

; TITLE OF INVENTION: Beta-Carotene Biosynthesis in

; TITLE OF INVENTION: Genetically Engineered Hosts

; NUMBER OF SEQUENCES: 79

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Amoco Corp., Patents and Licensing Dept

RESULT 15
US-08-096-043-6
; Sequence 6, Application US/08096043
; Patent No. 5530189
; GENERAL INFORMATION:
; APPLICANT: Ausich, Rodney L
; APPLICANT: Brinkhaus, Friedhelm L
; APPLICANT: Mukharji, Indrani
; APPLICANT: Proffitt, John H
; APPLICANT: Yarger, James G
; APPLICANT: Yen, Hwei-Che B
; TITLE OF INVENTION: Lycopene Biosynthesis in
; TITLE OF INVENTION: Genetically Engineered Hosts
; NUMBER OF SEQUENCES: 70
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amoco Corp., Patents and Licensing Dept
; STREET: 200 E Randolph St
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60680-0703
; COMPUTER READABLE FORM:

Search completed: November 12, 2003, 16:31:37
Job time : 27 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 12, 2003, 16:30:31 ; Search time 32 Seconds
(without alignments)

2082.460 Million cell updates/sec

Title: US-10-024-130A-6

Perfect score: 2051

Sequence: 1 MGSGLTMRYPDDIYPLKMK.....CRDAGVLQNRKSYNDKGQP 388

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 644079 seqs, 17179292 residues

Total number of hits satisfying chosen parameters: 644079

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications AA:*
- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
 - 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
 - 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 2051 | 100.0 | 388 | 12 | US-10-024-130A-6 |
| 2 | 2051 | 100.0 | 410 | 12 | US-10-024-130A-2 |
| 3 | 845 | 41.2 | 416 | 12 | US-10-205-194-67 |
| 4 | 820 | 40.0 | 417 | 10 | US-09-820-004-4 |
| 5 | 820 | 40.0 | 417 | 10 | US-09-820-004-5 |
| 6 | 820 | 40.0 | 417 | 10 | US-09-820-004-6 |
| 7 | 735 | 35.8 | 448 | 12 | US-10-032-585-7355 |
| 8 | 716.5 | 34.9 | 374 | 10 | US-09-820-004-2 |
| 9 | 675.5 | 32.9 | 520 | 15 | US-10-128-714-3505 |
| 10 | 675.5 | 32.9 | 520 | 15 | US-10-128-714-8505 |
| 11 | 227 | 11.1 | 363 | 12 | US-10-358-917-2 |
| 12 | 153 | 7.5 | 80 | 12 | US-10-024-130A-13 |
| 13 | 139 | 6.8 | 287 | 12 | US-10-358-917-10 |
| 14 | 131 | 6.4 | 314 | 12 | US-10-128-713A-16 |
| 15 | 117.5 | 5.7 | 316 | 15 | US-10-156-761-9192 |

| | | | | | | |
|----|-------|-----|------|----|---------------------|-------------------|
| 16 | 116 | 5.7 | 296 | 11 | US-09-941-947A-34 | Sequence 34, Appl |
| 17 | 116 | 5.7 | 296 | 12 | US-10-218-118-10 | Sequence 10, Appl |
| 18 | 112.5 | 5.5 | 440 | 10 | US-09-847-081B-2 | Sequence 2, Appl |
| 19 | 110.5 | 5.4 | 304 | 10 | US-09-738-626-4195 | Sequence 4195, Ap |
| 20 | 102.5 | 5.0 | 412 | 9 | US-09-371-307-76 | Sequence 76, Appl |
| 21 | 99.5 | 4.9 | 292 | 10 | US-09-738-626-6173 | Sequence 6173, Ap |
| 22 | 98 | 4.8 | 410 | 10 | US-09-847-081B-4 | Sequence 4, Appl |
| 23 | 96.5 | 4.7 | 316 | 10 | US-09-968-436B-4 | Sequence 4, Appl |
| 24 | 96.5 | 4.7 | 316 | 10 | US-09-968-436B-6 | Sequence 6, Appl |
| 25 | 96.5 | 4.7 | 581 | 9 | US-09-887-586A-48 | Sequence 48, Appl |
| 26 | 96.5 | 4.7 | 581 | 9 | US-09-903-012-48 | Sequence 48, Appl |
| 27 | 96.5 | 4.7 | 581 | 11 | US-09-900-797-48 | Sequence 48, Appl |
| 28 | 96.5 | 4.7 | 581 | 14 | US-10-041-007-18 | Sequence 18, Appl |
| 29 | 96.5 | 4.7 | 803 | 9 | US-09-759-010-7 | Sequence 7, Appl |
| 30 | 96.5 | 4.7 | 803 | 10 | US-09-968-436B-2 | Sequence 2, Appl |
| 31 | 96.5 | 4.7 | 803 | 15 | US-10-233-553-3 | Sequence 3, Appl |
| 32 | 94.5 | 4.6 | 759 | 10 | US-09-764-864-823 | Sequence 823, App |
| 33 | 94 | 4.6 | 342 | 15 | US-10-156-761-8563 | Sequence 8563, Ap |
| 34 | 93.5 | 4.6 | 1096 | 15 | US-10-128-714-3376 | Sequence 3376, Ap |
| 35 | 93.5 | 4.6 | 1096 | 15 | US-10-128-714-8376 | Sequence 8376, Ap |
| 36 | 93 | 4.5 | 682 | 9 | US-09-815-242-11214 | Sequence 11214, A |
| 37 | 92.5 | 4.5 | 945 | 12 | US-10-032-585-7114 | Sequence 7114, Ap |
| 38 | 92 | 4.5 | 344 | 11 | US-09-820-843A-35 | Sequence 35, Appl |
| 39 | 92 | 4.5 | 1572 | 15 | US-10-037-182-20 | Sequence 20, Appl |
| 40 | 92 | 4.5 | 1605 | 15 | US-10-037-182-18 | Sequence 18, Appl |
| 41 | 91.5 | 4.5 | 577 | 12 | US-10-025-145A-18 | Sequence 18, Appl |
| 42 | 91 | 4.4 | 1007 | 10 | US-09-843-676-86 | Sequence 86, Appl |
| 43 | 91 | 4.4 | 1007 | 10 | US-09-766-253-86 | Sequence 86, Appl |
| 44 | 91 | 4.4 | 1007 | 11 | US-09-438-486-86 | Sequence 86, Appl |
| 45 | 91 | 4.4 | 1007 | 15 | US-10-053-758-86 | Sequence 86, Appl |

ALIGNMENTS

RESULT 1

US-10-024-130A-6
; Sequence 6, Application US/100241130A
; Publication No. US20030157583A1
; GENERAL INFORMATION:
; APPLICANT: Stevens, Donna
; APPLICANT: Wang, Michelle
; APPLICANT: Rice, John
; APPLICANT: Lanning, Beth
; APPLICANT: Broadwell, David
; APPLICANT: Glassbrook, No. US20030157583Alman
; APPLICANT: Sevala, Veeresh
; APPLICANT: Crawford, John
; APPLICANT: Stewart, Sandy
; TITLE OF INVENTION: METHODS FOR DETERMINING SQUALENE
; TITLE OF INVENTION: SYNTHASE ACTIVITY
; FILE REFERENCE: 2148US
; CURRENT APPLICATION NUMBER: US/10/024,130A
; CURRENT FILING DATE: 2001-12-17
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 388
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-024-130A-6

| | | | | |
|-----------------------|-----------------|--|-----------|-------------|
| Query Match | 100.0% | Score 2051; | DB 12; | Length 388; |
| Best Local Similarity | 100.0% | Pred. No. 2.7e-199; | | |
| Matches 388; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |
| Qy | 1 | MGSGLTMRYPDDIYPLKMKRAIEKAEKQIPPEPHWGFCYSMLHKVSRFSVLVQQLNT | 60 | |
| Db | 1 | MGSGLTMRYPDDIYPLKMKRAIEKAEKQIPPEPHWGFCYSMLHKVSRFSVLVQQLNT | 60 | |
| Qy | 61 | ELRNVCFVLVLRALDTVEDDTSIPTDEKVPILIAFHRHIYDTDWHYSGCTKEYKILMD | 120 | |
| Db | 61 | ELRNVCFVLVLRALDTVEDDTSIPTDEKVPILIAFHRHIYDTDWHYSGCTKEYKILMD | 120 | |

```
QY 121 QFHVSAAFLELEKGYQEAIEITRRMGAGNAKFI COEVTETVDDYDEYCHYVAGLVGLG 180
Db 121 QFHVSAAFLELEKGYQEAIEITRRMGAGNAKFI COEVTETVDDYDEYCHYVAGLVGLG 180
QY 181 SKLFLAAGSEVLTPDWEAISNSMGLFLOKTNIIIRDYLEDEINEIPKSRMFPREIWKGYAD 240
Db 181 SKLFLAAGSEVLTPDWEAISNSMGLFLOKTNIIIRDYLEDEINEIPKSRMFPREIWKGYAD 240
QY 241 KLEDLKYEENTNKSVOCLNEMVTNALMHIEDCLKYMSVLRDPSIFRCAIPQIMAI GTLA 300
Db 241 KLEDLKYEENTNKSVOCLNEMVTNALMHIEDCLKYMSVLRDPSIFRCAIPQIMAI GTLA 300
QY 301 LCYNNEQVFRGVVKKLRRLGTAKVIDRTKTMADYVGFYDFSCMLKTKYVDKNDPNASKTLN 360
Db 301 LCYNNEQVFRGVVKKLRRLGTAKVIDRTKTMADYVGFYDFSCMLKTKYVDKNDPNASKTLN 360
QY 361 RLEAVOKLCRDAGVLQNRKSYVNDKGOP 388
Db 361 RLEAVOKLCRDAGVLQNRKSYVNDKGOP 388

RESULT 2
US-10-024-130A-2
; Sequence 2, Application US/10024130A
; Publication No. US20030157583A1
; GENERAL INFORMATION:
; APPLICANT: Stevens, Donna
; APPLICANT: Wang, Michelle
; APPLICANT: Rice, John
; APPLICANT: Lanning, Beth
; APPLICANT: Broadwell, David
; APPLICANT: Glassbrook, No. US20030157583A1man
; APPLICANT: Sevala, Veeresh
; APPLICANT: Crawford, John
; APPLICANT: Stewart, Sandy
; TITLE OF INVENTION: METHODS FOR DETERMINING SQUALENE
; TITLE OF INVENTION: SYNTHASE ACTIVITY
; FILE REFERENCE: 2148US
; CURRENT APPLICATION NUMBER: US/10/024,130A
; CURRENT FILING DATE: 2001-12-17
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 410
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-024-130A-2

Query Match 100.0%; Score 2051; DB 12; Length 410;
Best Local Similarity 100.0%; Pred. No. 2.9e-199;
Matches 388; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLGTMRLYPDDIYPLLLKMKRAIEKAEKQIPPEPHWGFCYSMLHKVSRFSLSVLIQOLNT 60
Db 1 MSLGTMRLYPDDIYPLLLKMKRAIEKAEKQIPPEPHWGFCYSMLHKVSRFSLSVLIQOLNT 60
QY 61 ELRNACVFLVLRALDTVEDDTSIPTDEKVPILIAFHRIHYDTDWHYSCGTKEYKILMD 120
Db 61 ELRNACVFLVLRALDTVEDDTSIPTDEKVPILIAFHRIHYDTDWHYSCGTKEYKILMD 120
QY 121 QFHVSAAFLELEKGYQEAIEITRRMGAGNAKFI COEVTETVDDYDEYCHYVAGLVGLG 180
Db 121 QFHVSAAFLELEKGYQEAIEITRRMGAGNAKFI COEVTETVDDYDEYCHYVAGLVGLG 180
QY 181 SKLFLAAGSEVLTPDWEAISNSMGLFLOKTNIIIRDYLEDEINEIPKSRMFPREIWKGYAD 240
Db 181 SKLFLAAGSEVLTPDWEAISNSMGLFLOKTNIIIRDYLEDEINEIPKSRMFPREIWKGYAD 240
QY 241 KLEDLKYEENTNKSVOCLNEMVTNALMHIEDCLKYMSVLRDPSIFRCAIPQIMAI GTLA 300
Db 241 KLEDLKYEENTNKSVOCLNEMVTNALMHIEDCLKYMSVLRDPSIFRCAIPQIMAI GTLA 300
```

```
QY 301 LCYNNEQVFRGVVKKLRRLGTAKVIDRTKTMADYVGFYDFSCMLKTKYVDKNDPNASKTLN 360
Db 301 LCYNNEQVFRGVVKKLRRLGTAKVIDRTKTMADYVGFYDFSCMLKTKYVDKNDPNASKTLN 360
QY 361 RLEAVOKLCRDAGVLQNRKSYVNDKGOP 388
Db 361 RLEAVOKLCRDAGVLQNRKSYVNDKGOP 388

RESULT 3
US-10-205-194-67
; Sequence 67, Application US/10205194
; Publication No. US20030134301A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; APPLICANT: Lee, Kevin
; APPLICANT: Dixon, Alistair
; APPLICANT: Brooksbank, Robert
; APPLICANT: Pinnoch, Robert
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
; FILE REFERENCE: WL-A-018201
; CURRENT APPLICATION NUMBER: US/10/205,194
; CURRENT FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: GB 0118354.0
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 177
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 67
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Squalene synthetase
US-10-205-194-67

Query Match 41.2%; Score 845; DB 12; Length 416;
Best Local Similarity 46.5%; Pred. No. 6.3e-77;
Matches 166; Conservative 71; Mismatches 108; Indels 12; Gaps 5;

QY 8 LRYPDIIYPLLLKMKRA-----IEKAEKQIPPEPHWGFCYSMLHKVSRFSLSVLIQOLNTEL 62
Db 7 LGHEEFYNLLRFRMGGRNFIPKMDRN-SLSNLSKCYKYQLDQTSRSPAIVQALDGD 65
QY 63 RNACVFLVLRALDTVEDDTSIPTDEKVPILIAFHRIHYDTDWHYSCGTKEYKILMDQF 122
Db 66 RHACVFLVLRALDTVEDDTSIPTDEKVPILIAFHRIHYDTDWHYSCGTKEYKILMDQF 125
QY 123 HHVSAAFLELEKGYQEAIEITRRMGAGNAKFI COEVTETVDDYDEYCHYVAGLVGLGSLK 182
Db 126 PTISLEFRNLAEKYQTIVADICHRMGCGMAEFLNKDVTQKQDWDKYCHYVAGLVGLGSLR 185
QY 183 LFLAAGSE--VLTPDWEAISNSMGLFLOKTNIIIRDYLEDEINEIPKSRMFPREIWKGYAD 240
Db 186 LFSASEFDPVIGDTEC-ANSMGLFLOKTNIIIRDYLEDQOE---GRQFMPQEWGKYK 241
QY 241 KLEDLKYEENTNKSVOCLNEMVTNALMHIEDCLKYMSVLRDPSIFRCAIPQIMAI GTLA 300
Db 242 KLEDLVKPEENVDAVAVKCLNELITNALQHPDVIYTLRLNRQSVNFCALPQWMAIATLA 301
QY 301 LCYNNEQVFRGVVKKLRRLGTAKVIDRTKTMADYVGFYDFSCMLKTKYVDKNDPNASK 357
Db 302 ACYNHGVFKGVVKIRKQAVTLMMDATNMPAVKAIYQYIEEIVHVRVPSNPSASK 358

RESULT 4
US-09-820-004-4
; Sequence 4, Application US/09820004
; Patent No. US20020142418A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
```



```
Db      :      :
363 ISTR 367

RESULT 7
US-10-032-585-7355
; Sequence 7355, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032.585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7355
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-032-585-7355

Query Match      35.8%; Score 735; DB 12; Length 448;
Best Local Similarity 40.8%; Pred. No. 1e-65;
Matches 160; Conservative 74; Mismatches 124; Indels 34; Gaps 8;

QY      1  MSLGTMRLRYDDI-----YPLLKMKRAIEAKKQIPPEPHMGFCYSMLHKVSR 50
Db      1  MGKFLQLLSHPTELKAVIQLFGRQPLHPGKRDVN--DKELVR-----CYELLNLT 52

QY      51  FSLVIOQLNTELRNVCVYLVLRALDVTEDTSTPTDEKVPILLAFHRIHYDTHWYSC 110
Db      53  FFAVIEELHPELDRAVMFYLVLRALDVTEDDTTSSIKIPLLREFDTKLTNKNWTFDG 112

QY      111  G--TKEYKILMDQFHVSAAFLKGLKEDLKYENTNK--SVQCLNEMVTNLMHIEDCLKYMV 162
Db      113  NGNENKDRVLVEFDKILNVYHLKPYQDIKSIIFKNGMGADYILDEBNVNGVATV 172

QY      163  DDYDEYCHYVAGLVGLSKLFLAA--GSEVITPDWEAISNMGFLFLOKTNIIIRDYLEDI 220
Db      173  EDVNLVCHYVAGLVGEGTLNLFVLANFGDKTLTENNAKADSMGLFLOKTNIIIRDYHEDL 232

QY      221  NEIPKSRMFWPREIWKGYADKLEDKYENTNK--SVQCLNEMVTNLMHIEDCLKYMV 277
Db      233  QD---GRSFWPREIWSKYTENLQDFHKVTPAKFAGVSCINELVNLALGHVTDCLDLYLS 289

QY      278  SLRDPSTFRCAIPQIMAIGTALCYNNEQVFRGVVGLRGLTAKVIDRTKTMADYVYAF 337
Db      290  LVKDPSSFSFCAIPQMAVATLAEVYNNPKVLHGVVVKIRKGTTCRLILESRTPGKVIF 349

QY      338  YDFSCMLKTKVDKNDPNASKTNLRLEAVQKLC 369
Db      350  KEYIQVINHKSSVRDPNLYLKIGIKGEBIEQYC 381

RESULT 8
US-09-820-004-2
; Sequence 2, Application US/09820004
; Patent No. US20020142418A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001201
; CURRENT APPLICATION NUMBER: US/09/820.004
; CURRENT FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2

Query Match      32.9%; Score 675.5; DB 15; Length 520;
```

| | | | | | |
|--|-----|---|-----|--|--|
| | | | | | |
| <hr/> | | | | | |
| Best Local Similarity 43.7%; Pred. No. 1.4e-59; | | | | | |
| Matches 143; Conservative 59; Mismatches 108; Indels 17; Gaps 7; | | | | | |
| <hr/> | | | | | |
| Qy | 40 | CYSMLHKVRSFSLSVIOQLNTELRNAVCFYLVLRALDVTDEDDTSIPTDEKVPILIAFHR | 99 | | |
| Dd | 37 | CFKFELDTSRSFSAVIKEHPHELLPNCIFYLTLRGLDTIEDTSPLETKPELLRGFKD | 96 | | |
| Qy | 100 | HIYDTHWHYSCGKTKEY--ILMDQHVVHVSAAFLELEKGYOEAIEETRRMGAGMAKFICQ | 157 | | |
| Dd | 97 | VLEGDGWNFTGNRPPEEKRELLVFQHNIVITFEFKRLKPAYKAVIKDITEKMNGMADY-CR | 155 | | |
| Qy | 158 | E-----VSTVDYDYCHYVAGVLGLSKFLAA--CSEVLTPDWAEISNMGLFIQ | 208 | | |
| Dd | 156 | KAADDDASVKTYEYDIYCYYAGLVGGESTRLFVEAFEGNPALLKRTE-LYKSMLGIQ | 214 | | |
| Qy | 209 | KTNIIRDYLEIDINEIPKSRMFWEIRGWKYADKLDKYEENTKNKSVOCLNEVMVTNALMH | 268 | | |
| Dd | 215 | KTNIIRDVRDFDD--GROFWPKETWSKVHTWFEDLFKENREAALNCSEMWLNALRH | 271 | | |
| Qy | 269 | IDECLKMYSLRDPSIFRFCAIQIMAIGTLALCYNNEQVFRGVVKLRGLGTAK-VIDRT | 327 | | |
| Dd | 272 | AEECLFYLAGLREQSVNFCAIPQSMAIATLSLCFRNPAIFERNIKIKGEACOLMMWEST | 331 | | |
| Qy | 328 | KTWADVGYAfydfscmlktkvkndpn | 354 | | |
| Dd | 332 | ONLRILEAFRRYAREIHKKNTPKDPN | 358 | | |
| <hr/> | | | | | |
| RESULT 10 | | | | | |
| US-10-128-714-8505 | | | | | |
| ; Sequence 8505, Application US/10128714 | | | | | |
| ; Publication No. US20030119013A1 | | | | | |
| ; GENERAL INFORMATION: | | | | | |
| ; APPLICANT: Jiang, Bo | | | | | |
| ; APPLICANT: Hu, Wengi | | | | | |
| ; APPLICANT: Tishkoff, Daniel | | | | | |
| ; APPLICANT: Zamudio, Carlos | | | | | |
| ; APPLICANT: Eroshkin, Alexey M | | | | | |
| ; APPLICANT: Lemieux, Sebastien M | | | | | |
| ; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and | | | | | |
| ; FILE REFERENCE: Methods of Use | | | | | |
| ; CURRENT APPLICATION NUMBER: US/10/128,714 | | | | | |
| ; CURRENT FILING DATE: 2002-04-23 | | | | | |
| ; PRIOR APPLICATION NUMBER: US 60/285,697 | | | | | |
| ; PRIOR FILING DATE: 2001-04-23 | | | | | |
| ; PRIOR APPLICATION NUMBER: US 60/287,066 | | | | | |
| ; PRIOR FILING DATE: 2001-04-27 | | | | | |
| ; PRIOR APPLICATION NUMBER: US 60/295,890 | | | | | |
| ; PRIOR FILING DATE: 2001-06-05 | | | | | |
| ; PRIOR APPLICATION NUMBER: US 60/303,899 | | | | | |
| ; PRIOR FILING DATE: 2001-07-09 | | | | | |
| ; PRIOR APPLICATION NUMBER: US 60/316,362 | | | | | |
| ; PRIOR FILING DATE: 2001-08-31 | | | | | |
| ; NUMBER OF SEQ ID NOS: 8603 | | | | | |
| ; SOFTWARE: PatentIn version 3.1 | | | | | |
| ; SEQ ID NO 8505 | | | | | |
| ; LENGTH: 528 | | | | | |
| ; TYPE: PRT | | | | | |
| ; ORGANISM: Aspergillus fumigatus | | | | | |
| US-10-128-714-8505 | | | | | |
| <hr/> | | | | | |
| Query Match 32.9%; Score 675.5; DB 15; Length 528; | | | | | |
| Best Local Similarity 43.7%; Pred. No. 1.4e-59; | | | | | |
| Matches 143; Conservative 59; Mismatches 108; Indels 17; Gaps 7; | | | | | |
| <hr/> | | | | | |
| Qy | 40 | CYSMLHKVRSFSLSVIOQLNTELRNAVCFYLVLRALDVTDEDDTSIPTDEKVPILIAFHR | 99 | | |
| Dd | 45 | CFKFELDTSRSFSAVIKEHPHELLPNCIFYLTLRGLDTIEDTSPLETKPELLRGFKD | 104 | | |
| Qy | 100 | HIYDTHWHYSCGKTKEY--ILMDQHVVHVSAAFLELEKGYOEAIEETRRMGAGMAKFICQ | 157 | | |
| Dd | 105 | VLEGDGWNFTGNRPPEEKRELLVFQHNIVITFEFKRLKPAYKAVIKDITEKMNGMADY-CR | 163 | | |

| QY | 255 | VQCLNEMVTNALMHIEDCLKYMWVSLRD | 281 |
|--|------------------|---|--------------------|
| Db | 207 | -----HYIDLWEYAAIAE | 220 |
| RESULT 14 | | | |
| US-10-128-713A-16 | | | |
| Sequence 16, Application US/10128713A | | | |
| Publication No. US20030170847A1 | | | |
| GENERAL INFORMATION: | | | |
| APPLICANT: Bramucci, Michael G | | | |
| TITLE OF INVENTION: Genes Involved in Isoprenoid Compound Production | | | |
| FILE REFERENCE: CL-1788 | | | |
| CURRENT APPLICATION NUMBER: US/10/128,713A | | | |
| CURRENT FILING DATE: 2002-04-22 | | | |
| NUMBER OF SEQ ID NOS: 36 | | | |
| SOFTWARE: Microsoft Office 97 | | | |
| SEQ ID NO 16 | | | |
| LENGTH: 314 | | | |
| TYPE: PRT | | | |
| ORGANISM: Rhodococcus erythropolis | | | |
| US-10-128-713A-16 | | | |
| Query Match | 6.4% | Score 131; | DB 12; Length 314; |
| Best Local Similarity | 24.4%; | Pred. No. 0.00011; | |
| Matches | 58; Conservative | 40; Mismatches 118; | Indels 22; Gaps 7; |
| QY | 39 | FCYSMLHKVSRFSLSVLTQNLTELNRNAVCFVYLVLRALDVTVEDTSTPTDEKVPILIAFH | 98 |
| Db | 10 | FCEDVTREHGRTYFLATLLEPPERRAVHALYAFARVVDVDEPSPGHERGTVLADVER | 69 |
| QY | 99 | RHIYDVTDMHYSCTGKEYKILMDQPHVSAAF-----LELEKGYQEAIEIITRMGAMA | 152 |
| Db | 70 | AAVTALDNPTATGGPPSTPLD-LTRVLPAPADAVKTFDIPRAYFDFAFFESMRMDAPDTA | 128 |
| QY | 153 | KFICQEVETVDDYDEYCHYVAGLVGLSKFLAAGSEVLTPPOWEAL--SNSMGLFLQKT | 210 |
| Db | 129 | KF-RPVYNTMDLAEYMGSAVIGLQMLPIL-----GVSPQCEAVVPASNGLGEAFQLT | 182 |
| QY | 211 | NIIRDYLEDINEIPKSRMFWPREIWGKYADKLEDLYEENTNKSVCQCLNEWVTNALMH | 268 |
| Db | 183 | NFIRDVGEDLD---RGRLYLPAFGFAAFGVDIEMLEHGRRTG----TVDVRVKRALAH | 233 |
| RESULT 15 | | | |
| US-10-156-761-9192 | | | |
| Sequence 9192, Application US/10156761 | | | |
| Publication No. US20030119018A1 | | | |
| GENERAL INFORMATION: | | | |
| APPLICANT: OMURA, SATOSHI | | | |
| APPLICANT: IKEDA, HARUO | | | |
| APPLICANT: ISHIKAWA, JUN | | | |
| APPLICANT: HORIKAWA, HIROSHI | | | |
| APPLICANT: SHIBA, TADAYOSHI | | | |
| APPLICANT: SAKAKI, YOSHIYUKI | | | |
| APPLICANT: HATTORI, MASAHIRA | | | |
| TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES | | | |
| FILE REFERENCE: 249-262 | | | |
| CURRENT APPLICATION NUMBER: US/10/156,761 | | | |
| CURRENT FILING DATE: 2002-05-29 | | | |
| PRIOR APPLICATION NUMBER: JP 2001-204089 | | | |
| PRIOR FILING DATE: 2001-05-30 | | | |
| PRIOR APPLICATION NUMBER: JP 2001-272697 | | | |
| PRIOR FILING DATE: 2001-08-02 | | | |
| NUMBER OF SEQ ID NOS: 15109 | | | |
| SEQ ID NO 9192 | | | |
| LENGTH: 316 | | | |
| TYPE: PRT | | | |
| ORGANISM: Streptomyces avermitilis | | | |
| US-10-156-761-9192 | | | |
| Query Match | 5.7% | Score 117.5; | DB 15; Length 316; |

Best Local Similarity 21.9%; Pred. No. 0.0026;
Matches 54; Conservative 33; Mismatches 89; Indels 71; Gaps 8;

| | | | |
|----|-----|---|-----|
| Qy | 22 | RAIEKAEK-QIPPEPHWGCYSMLHKVSRFSLVIOQLNTELRNAVCFYLVLRALDTVE | 80 |
| Db | 3 | RTVESEQHVSAPVLAAYSCEYVTGQARNFAYGIRLLPTPKRAMSAYAFSRRVDDIG | 62 |
| Qy | 81 | D-----DTSIPTDEKVPILIAFHRHIYDTDMHYSCTKEYKI | 117 |
| Db | 63 | DGALAPDVKAARLEDTALLSRVRDGRVDEDDTDPVAVA----- | 101 |
| Qy | 118 | LMDQPHHVSAAFLLEKGYOEAIIEITRRMGAGMAKFCOEVEVDDYDEYCHYVAGLVG | 177 |
| Db | 102 | -----LAHAASQFPFPLGGLDELIDGVLMDVRG-----ETYETWDDLKYCRCVAGAIG | 150 |
| Qy | 178 | ---LGLSKLFLAAGSEVLTPDWEAIS---NSMGLFLQKTNIIIRDYLEDINEIPKSRMFWP | 231 |
| Db | 151 | RVSLGV-----FGTEPGARGAERASEYADTLGLALQLTNILRDVREDAE---GGRTYLP | 201 |
| Qy | 232 | REIWGY | 238 |
| Db | 202 | ADDLAKF | 208 |

Search completed: November 12, 2003, 16:36:12
Job time : 42 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 12, 2003, 16:27:10 ; Search time 21 Seconds
(without alignments)
1776.831 Million cell updates/sec

Title: US-10-024-130A-6

Perfect score: 2051

Sequence: 1 MGSGLTMLRYDDIYPLLMK.....CRDAGVLQNRKSYVNDKGQP 388

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_76:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|-------------|
| 1 | 2051 | 100.0 | 410 | 2 | S54251 |
| 2 | 1702.5 | 83.0 | 413 | 2 | T05262 |
| 3 | 1697 | 82.7 | 413 | 2 | T44924 |
| 4 | 1692 | 82.5 | 413 | 2 | H85408 |
| 5 | 1671 | 81.5 | 413 | 2 | JC5030 |
| 6 | 1659 | 80.9 | 412 | 2 | JC5031 |
| 7 | 1617 | 78.8 | 411 | 2 | S71771 |
| 8 | 1605.5 | 78.3 | 412 | 2 | T05285 |
| 9 | 1446 | 70.5 | 401 | 2 | T00489 |
| 10 | 1437.5 | 70.1 | 376 | 2 | T02167 |
| 11 | 848 | 41.3 | 416 | 2 | S52075 |
| 12 | 845 | 41.2 | 416 | 2 | A45105 |
| 13 | 820 | 40.0 | 417 | 2 | A45998 |
| 14 | 820 | 40.0 | 417 | 2 | I38245 |
| 15 | 814 | 39.7 | 417 | 2 | I52090 |
| 16 | 721 | 35.2 | 444 | 1 | S46682 |
| 17 | 703.5 | 34.3 | 460 | 2 | B48057 |
| 18 | 700.5 | 34.2 | 460 | 2 | T40581 |
| 19 | 278.5 | 13.6 | 390 | 2 | A84226 |
| 20 | 208 | 10.1 | 362 | 2 | T45141 |
| 21 | 181 | 8.8 | 287 | 2 | A92031 |
| 22 | 175.5 | 8.6 | 277 | 2 | S75973 |
| 23 | 145.5 | 7.1 | 290 | 2 | A81118 |
| 24 | 139 | 6.8 | 287 | 2 | C90061 |
| 25 | 134.5 | 6.6 | 290 | 2 | H81902 |
| 26 | 122.5 | 6.0 | 302 | 2 | S70735 |
| 27 | 122 | 5.9 | 312 | 2 | T35400 |
| 28 | 121.5 | 5.9 | 325 | 2 | D75466 |
| 29 | 118.5 | 5.8 | 336 | 2 | S32170 |

| | | | | | | |
|----|-------|-----|------|---|--------|--------------------|
| 30 | 118 | 5.8 | 310 | 2 | AC2035 | phytoene synthase |
| 31 | 118 | 5.8 | 410 | 2 | S68307 | phytoene synthase |
| 32 | 115.5 | 5.6 | 296 | 2 | S52587 | prephytoene pyroph |
| 33 | 115 | 5.6 | 296 | 2 | E37802 | crbB protein - Erw |
| 34 | 113 | 5.5 | 222 | 2 | E84320 | phytoene synthase |
| 35 | 111 | 5.4 | 2166 | 2 | G70163 | hypothetical prote |
| 36 | 109.5 | 5.3 | 339 | 2 | S44403 | geranylgeranyl-dip |
| 37 | 109.5 | 5.3 | 674 | 2 | S46092 | probable membrane |
| 38 | 107.5 | 5.2 | 309 | 2 | B39273 | geranylgeranyl-dip |
| 39 | 107.5 | 5.2 | 734 | 2 | G86171 | hypothetical prote |
| 40 | 107 | 5.2 | 337 | 2 | S45360 | phytoene synthase |
| 41 | 105.5 | 5.1 | 234 | 2 | A55548 | crbM protein - Sta |
| 42 | 104 | 5.1 | 322 | 2 | H84299 | phytoene synthase |
| 43 | 104 | 5.1 | 423 | 2 | S54135 | phytoene synthase |
| 44 | 104 | 5.1 | 473 | 2 | H85482 | hypothetical prote |
| 45 | 104 | 5.1 | 473 | 2 | A90632 | hypothetical prote |

ALIGNMENTS

RESULT 1

S54251

farnesyl-diphosphate farnesyltransferase (EC 2.5.1.21) T4L20.220 - Arabidopsis thaliana
N;Alternate names: protein T4L20.220; squalene synthase

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 15-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 23-Jul-1999

C;Accession: S54251, T05284

R;Kribbi, R.; Boronati, A.; Karst, F.

(submitted to the EMBL Data Library, April 1995

A;Description: Molecular cloning, expression and characterization of cDNAs for Arabidopsis

A;Reference number: S54251

A;Accession: S54251

A;Molecule type: mRNA

A;Residues: 1-410 <KRI>

A;Cross-references: EMBL:X86692; NID:g798819; PIDN:CAA60385.1; PID:g798820

R;Bevan, M.; Terry, N.; Ardiles, W.; Buysahaert, C.; Dasseville, R.; De Clerck, R.; De

ewes, H.W.; Mayer, K.F.X.; Schueller, C.

submitted to the Protein Sequence Database, September 1998

A;Reference number: Z15406

A;Accession: T05284

A;Molecule type: DNA

A;Residues: 1-410 <BEV>

A;Cross-references: EMBL:AL023094

A;Experimental source: cultivar Columbia; BAC clone T4L20

C;Genetics:

A;Map position: 4

A;Introns: 65/3; 80/1; 110/1; 135/2; 158/3; 208/3; 243/3; 292/3; 318/1; 347/3; 378/2; 392/3

C;Note: T4L20.220

C;Superfamily: farnesyl-diphosphate farnesyltransferase

C;Keywords: transferase

Query Match 100.0%; Score 2051; DB 2; Length 410;

Best Local Similarity 100.0%; Pred. No. 2.1e-152;

Matches 388; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGSGLTMLRYDDIYPLLMKRAIERAKQIPEPHWGCYCSMLHKVSRFSLSVIOQLNT 60

Db 1 MGSGLTMLRYDDIYPLLMKRAIERAKQIPEPHWGCYCSMLHKVSRFSLSVIOQLNT 60

Qy 61 ELRNACVFLVLRALDTVEDDTSIPTDEKVPILIAFHRHIYDTDWHYSGTKKEYKILMD 120

Db 61 ELRNACVFLVLRALDTVEDDTSIPTDEKVPILIAFHRHIYDTDWHYSGTKKEYKILMD 120

Qy 121 QFHVSAAFLELEKGYQEAIEEITRRMGAGMAKFIQCVETVDDYDEYCHYVAGLVGL 180

Db 121 QFHVSAAFLELEKGYQEAIEEITRRMGAGMAKFIQCVETVDDYDEYCHYVAGLVGL 180

Qy 181 SKLFLAAGSEVLTPDWEAISNSMGLFLOKTNIIIRDYLEDINEIPKSRMFWPRETWKYAD 240

Db 181 SKLFLAAGSEVLTPDWEAISNSMGLFLOKTNIIIRDYLEDINEIPKSRMFWPRETWKYAD 240

Qy 241 KLEDKYEENTNKSQVCLNEMVTNMLHIEDCLKYMVSLRDPISIFRCAIPQIMAIGTLA 300

Db 241 KLEDLKYEENTNKSVQCLNEMVTNALMHIHEDCLKYMVSLRDPISFRFCAIPQIMAIGTLA 300
QY 301 LCYNNEOVFGVVKLRGLTAKVIDRTKTWADVYGAFYDFSCMLKTKVDKNDPNASKTLN 360
Db 301 LCYNNEOVFGVVKLRGLTAKVIDRTKTWADVYGAFYDFSCMLKTKVDKNDPNASKTLN 360
QY 361 RLEAVQKLCRDAGVQLQRKSYVNDKGQP 388
Db 361 RLEAVQKLCRDAGVQLQRKSYVNDKGQP 388
RESULT 2
T05262
farnesyl-diphosphate farnesyltransferase (EC 2.5.1.21) - soybean
N:Alternate names: squalene synthase
C:Species: Glycine max (soybean)
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 20-Jun-2000
C:Accession: T05262
R:Hata, S.
submitted to the EMBL Data Library, September 1997
A:Description: CDNA cloning of squalene synthase genes from mono- and dicotyledonous plants
A:Reference number: Z14155
A:Accession: T05262
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-413 <HAT>
A:Cross-references: EMBL:AB007503; PIDN:BAA22559.1
A:Experimental source: cultivar Akisengoku
C:Function:
A:Description: catalyzes conversion of two molecules of farnesyl diphosphate into squalene
C:Superfamily: farnesyl-diphosphate farnesyltransferase
C:Keywords: transferase
Query Match 83.0%; Score 1702.5; DB 2; Length 413;
Best Local Similarity 82.4%; Pred. No. 3.4e-125;
Matches 318; Conservative 34; Mismatches 31; Indels 3; Gaps 2;
QY 1 MGSGLTMLRYDDIYPLLLKMKRAEKAQKQIPPEPHMGFCYSMLHKVSRFSFLVIQQLNT 60
Db 1 MGSGLTMLRYDDIYPLLLKMKRAEKAQKQIPPEPHMGFCYSMLHKVSRFSFLVIQQLGI 60
QY 61 ELRNAVCVFLVLRALDTVEDDTSIPTDEKVPILIAFHRHIYDTHWYSCGTKEYKILMD 120
Db 61 ELRNAVCVFLVLRALDTVEDDTSIPTDEKVPILIAFHRHIYDTHWYSCGTKEYKILMD 120
QY 121 QPHHVSAAFLKLEKGYQEAIEITRRMGAGMAKFCQEVETVDDYDEYCHYVAGLVGL 180
Db 121 QPHHVSAAFLKLEKGYQEAIEITRRMGAGMAKFCQEVETVDDYDEYCHYVAGLVGL 180
QY 181 SKLFLAAGSEVLTDPWEAISNSMGLFLOKTNIIIRDYLEDINEIPKSRMFWPREIWKGYD 240
Db 181 SKLFLAAGSEVLTDPWEAISNSMGLFLOKTNIIIRDYLEDINEIPKSRMFWPREIWKGYD 240
QY 241 KLEDLKYEENTNKSVQCLNEMVTNALMHIHEDCLKYMVSLRDPISFRFCAIPQIMAIGTLA 300
Db 241 KLEDLKYEENTNKSVQCLNEMVTNALMHIHEDCLKYMVSLRDPISFRFCAIPQIMAIGTLA 300
QY 301 LCYNNEOVFGVVKLRGLTAKVIDRTKTWADVYGAFYDFSCMLKTKVDKNDPNASKTLN 360
Db 301 LCYNNEOVFGVVKLRGLTAKVIDRTKTWADVYGAFYDFSCMLKTKVDKNDPNASKTLN 360
QY 361 RLEAVQKLCRDAGVQLQRKSY-VNDK 385
Db 361 RLEAVQKLCRDAGVQLQRKSY-VNDE 384
RESULT 3
T44924
farnesyl-diphosphate farnesyltransferase (EC 2.5.1.21) 2 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 18-Feb-2000
C:Accession: T44924

R:Kribbi, R.; Arro, M.; del Arco, A.; Gonzalez, V.; Balcells, L.L.; Delourme, D.; Ferrer
submitted to the EMBL Data Library, May 1997
A:Description: Cloning and characterization of the Arabidopsis thaliana SQS1 gene encoding
the sterol pathway.
A:Reference number: Z22869
A:Accession: T44924
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-413 <KRL>
A:Cross-references: EMBL:AF004396; PIDN:AAB61927.1
C:Genetics:
A:Gene: SQS2
A:Introns: 65/3; 80/1; 110/1; 135/2; 158/3; 208/3; 243/3; 292/3; 318/1; 347/3; 378/2; 39
C:Superfamily: farnesyl-diphosphate farnesyltransferase
C:Keywords: transferase
Query Match 82.7%; Score 1697; DB 2; Length 413;
Best Local Similarity 80.6%; Pred. No. 9.3e-125;
Matches 312; Conservative 38; Mismatches 37; Indels 0; Gaps 0;
QY 1 MGSGLTMLRYDDIYPLLLKMKRAEKAQKQIPPEPHMGFCYSMLHKVSRFSFLVIQQLNT 60
Db 1 MGSGLTMLRYDDIYPLLLKMKRAEKAQKQIPPEPHMGFCYSMLHKVSRFSFLVIQQLGT 60
QY 61 ELRNAVCVFLVLRALDTVEDDTSIPTDEKVPILIAFHRHIYDTHWYSCGTKEYKILMD 120
Db 61 ELRNAVCVFLVLRALDTVEDDTSIPTDEKVPILIAFHRHIYDTHWYSCGTKEYKILMD 120
QY 121 QPHHVSAAFLKLEKGYQEAIEITRRMGAGMAKFCQEVETVDDYDEYCHYVAGLVGL 180
Db 121 QPHHVSAAFLKLEKGYQEAIEITRRMGAGMAKFCQEVETVDDYDEYCHYVAGLVGL 180
QY 181 SKLFLAAGSEVLTDPWEAISNSMGLFLOKTNIIIRDYLEDINEIPKSRMFWPREIWKGYD 240
Db 181 SKLFLAAGSEVLTDPWEAISNSMGLFLOKTNIIIRDYLEDINEIPKSRMFWPREIWKGYD 240
QY 241 KLEDLKYEENTNKSVQCLNEMVTNALMHIHEDCLKYMVSLRDPISFRFCAIPQIMAIGTLA 300
Db 241 KLEDLKYEENTNKSVQCLNEMVTNALMHIHEDCLKYMVSLRDPISFRFCAIPQIMAIGTLA 300
QY 301 LCYNNEOVFGVVKLRGLTAKVIDRTKTWADVYGAFYDFSCMLKTKVDKNDPNASKTLN 360
Db 301 LCYNNEOVFGVVKLRGLTAKVIDRTKTWADVYGAFYDFSCMLKTKVDKNDPNASKTLN 360
QY 361 RLEAVQKLCRDAGVQLQRKSYVNDKGQ 387
Db 361 RLEAVQKLCRDAGVQLQRKSYVNDKGQ 387
RESULT 4
E85408
probable squalene synthase [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 02-Mar-2001
C:Accession: E85408
R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488; PMID:10617198
A:Accession: E85408
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-413 <STO>
A:Cross-references: GB:NC_001268; NID:g7270415; PIDN:CAB80182.1; GSPDB:GN00140
C:Genetics:
A:Gene: AT4g34650
A:Map position: 4
C:Superfamily: farnesyl-diphosphate farnesyltransferase
Query Match 82.5%; Score 1692; DB 2; Length 413;
Best Local Similarity 80.6%; Pred. No. 2.3e-124;
Matches 312; Conservative 37; Mismatches 38; Indels 0; Gaps 0;


```
QY 1 MGSIGTMLRPDDIYPLLLKMKRAIEKAEKQIPPEPHWGFYCSMLHKVSRFSLSLVIQOLNT 60
DB 1 MGSLSITLIRHPDELYPLLLKMLKLAITKAQKQIPLEPHLAFYCSILHKVSRFSLSLVIQOLGT 60
QY 61 ELRNACVFLVLRALDVTDEDDTSIPTDEKVPILIAFHRHIYDTHYSGTKEYKILMD 120
DB 61 ELRNACVFLVLRALDVTDEDDTSIPTDEKVPILIAFHRHIYDTHYSGTKEYKILMD 120
QY 121 QFHHVSAFLELEKGYQEAIEETTRMGAGMAKFI COEVTVDYDYCHYVAGLVGLG 180
DB 121 QFHHVSAFLELEKGYQEAIEDITKMGAGMAKFI KEVETIDYDYCHYVAGLVGLG 180
QY 181 SKLFLAAGSEVLTPDWEAISNSMGLFLOKTNIIIRDYLEINEIPKSRMFWPREIWKYAD 240
DB 181 SKLFLAAGSEVLTPDWEAISNSMGLFLOKTNIIIRDYLEINEIPKSRMFWPREIWKYAD 240
QY 241 KLEDLKYEENTNKSVOCLNEMVTNALMHIEDCLKYMVSLRDPSPISFRFCAIPQIMAGTGLA 300
DB 241 KLEDLKYEENTNKSVOCLNEMVTNALMHIEDCLKYMVSLRDPSPISFRFCAIPQIMAGTGLA 300
QY 301 LCYNNEQVFRGVVKKLRGLTAKVIDRTKTMDVYGFYDFSCMLKTKVDKNDPNASKTLN 360
DB 301 LCYNNEQVFRGVVKKLRGLTAKVIDRTKTMDVYGFYDFSCMLKTKVDKNDPNASKTLN 360
QY 361 RLEAVOKLDRDAGVLQNRKSYVNDKQ 387
DB 361 RLETIKFKCKENGHLKRRKSYVNDKQ 387

RESULT 5
JC5030
farnesyl-diphosphate farnesyltransferase (EC 2.5.1.21) 1 - Glycyrhrhiza glabra L.
N;Alternate names: farnesyltransferase; squalene synthase
C;Species: Glycyrhrhiza glabra L.
C;Date: 21-Jan-1997 #sequence_revision 21-Jan-1997 #text_change 20-Jun-2000
C;Accession: JC5030
R;Hayashi, H.; Hiraoka, N.; Ikeshiro, Y.
Biol. Pharm. Bull. 19, 1387-1389, 1996
A;Title: Molecular cloning and functional expression of cDNAs for Glycyrhrhiza glabra squ
A;Reference number: JC5030; MUID:97070595; PMID:8913521
A;Accession: JC5030
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-412 <HAY>
A;Cross-references: DDBJ:D86409; NID:g1449162; PIDN:BAA13083.1; PID:g1449163
C;Genetics:
A;Gene: G5QS1
C;Function:
C;Description: catalyzes the conversion of two molecules of farnesyl diphosphate into sq
C;Superfamily: farnesyl-diphosphate farnesyltransferase
C;Keywords: transferase

Query Match 81.5%; Score 1671; DB 2; Length 413;
Best Local Similarity 81.4%; Pred. No. 9.9e-123;
Matches 311; Conservative 34; Mismatches 35; Indels 2; Gaps 1;

QY 1 MGSIGTMLRPDDIYPLLLKMKRAIEKAEKQIPPEPHWGFYCSMLHKVSRFSLSLVIQOLNT 60
DB 1 MGSIGTMLRPDDIYPLLLKMKRAIEKAEKQIPPEPHWGFYCSMLHKVSRFSLSLVIQOLNT 60
QY 61 ELRNACVFLVLRALDVTDEDDTSIPTDEKVPILIAFHRHIYDTHYSGTKEYKILMD 120
DB 61 ELRNACVFLVLRALDVTDEDDTSIPTDEKVPILIAFHRHIYDTHYSGTKEYKILMD 120
QY 121 QFHHVSAFLELEKGYQEAIEETTRMGAGMAKFI COEVTVDYDYCHYVAGLVGLG 180
DB 121 QFHHVSAFLELEKGYQEAIEDITKMGAGMAKFI KEVETVDYDYCHYVAGLVGLG 180
QY 181 SKLFLAAGSEVLTPDWEAISNSMGLFLOKTNIIIRDYLEINEIPKSRMFWPREIWKYAD 240
DB 181 SKLFLAAGSEVLTPDWEAISNSMGLFLOKTNIIIRDYLEINEIPKSRMFWPREIWKYAD 240
QY 241 KLEDLKYEENTNKSVOCLNEMVTNALMHIEDCLKYMVSLRDPSPISFRFCAIPQIMAGTGLA 300
DB 241 KLEDLKYEENTNKSVOCLNEMVTNALMHIEDCLKYMVSLRDPSPISFRFCAIPQIMAGTGLA 300
```

```
DB 239 KLEDLKYEENSVAQVCLNDVMVTNALLHAEDCLKYMSALRDMISIFRFCAIPQIMAGTGLA 298
QY 301 LCYNNEQVFRGVVKKLRGLTAKVIDRTKTMDVYGFYDFSCMLKTKVDKNDPNASKTLN 360
DB 299 LCYNNEQVFRGVVKKLRGLTAKVIDRTKTMDVYGFYDFSCMLKTKVDKNDPNASKTLN 358
QY 361 RLEAVOKLDRDAGVLQNRKSYV 382
DB 359 RLEAIQKTCRESGLLSKRKPYI 380

RESULT 6
JC5031
farnesyl-diphosphate farnesyltransferase (EC 2.5.1.21) 2 - Glycyrhrhiza glabra L.
N;Alternate names: farnesyltransferase; squalene synthase
C;Species: Glycyrhrhiza glabra L.
C;Date: 21-Jan-1997 #sequence_revision 21-Jan-1997 #text_change 20-Jun-2000
C;Accession: JC5031
R;Hayashi, H.; Hiraoka, N.; Ikeshiro, Y.
Biol. Pharm. Bull. 19, 1387-1389, 1996
A;Title: Molecular cloning and functional expression of cDNAs for Glycyrhrhiza glabra squ
A;Reference number: JC5030; MUID:97070595; PMID:8913521
A;Accession: JC5031
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-412 <HAY>
A;Cross-references: DDBJ:D86410; NID:g1449164; PIDN:BAA13084.1; PID:g1449165
C;Genetics:
A;Gene: G5QS2
C;Function:
C;Description: catalyzes the conversion of two molecules of farnesyl diphosphate into sq
C;Superfamily: farnesyl-diphosphate farnesyltransferase
C;Keywords: transferase

Query Match 80.9%; Score 1659; DB 2; Length 412;
Best Local Similarity 79.9%; Pred. No. 8.5e-122;
Matches 307; Conservative 42; Mismatches 33; Indels 2; Gaps 1;

QY 1 MGSIGTMLRPDDIYPLLLKMKRAIEKAEKQIPPEPHWGFYCSMLHKVSRFSLSLVIQOLNT 60
DB 1 MGSIGTMLRPDDIYPLLLKMKRAIEKAEKQIPPEPHWGFYCSMLHKVSRFSLSLVIQOLNT 60
QY 61 ELRNACVFLVLRALDVTDEDDTSIPTDEKVPILIAFHRHIYDTHYSGTKEYKILMD 120
DB 61 ELRNACVFLVLRALDVTDEDDTSIPTDEKVPILIAFHRHIYDTHYSGTKEYKILMD 120
QY 121 QFHHVSAFLELEKGYQEAIEETTRMGAGMAKFI COEVTVDYDYCHYVAGLVGLG 180
DB 121 QFHHVSAFLELEKGYQEAIEDITKMGAGMAKFI KEVETVDYDYCHYVAGLVGLG 180
QY 181 SKLFLAAGSEVLTPDWEAISNSMGLFLOKTNIIIRDYLEINEIPKSRMFWPREIWKYAD 240
DB 181 SKLFLAAGSEVLTPDWEAISNSMGLFLOKTNIIIRDYLEINEIPKSRMFWPREIWKYAD 240
QY 241 KLEDLKYEENTNKSVOCLNEMVTNALMHIEDCLKYMVSLRDPSPISFRFCAIPQIMAGTGLA 300
DB 239 KLEDLKYEENSVAQVCLNDVMVTNALLHAEDCLKYMAALRDLAIFRFCAIPQIMAGTGLA 298
QY 301 LCYNNEQVFRGVVKKLRGLTAKVIDRTKTMDVYGFYDFSCMLKTKVDKNDPNASKTLN 360
DB 299 LCYNNEQVFRGVVKKLRGLTAKVIDRTKTMDVYGFYDFSCMLKTKVDKNDPNASKTLN 358
QY 361 RLEAVOKLDRDAGVLQNRKSYVND 384
DB 359 RLEAIQKTCRESGLLNKRKSLLRN 382

RESULT 7
S71771
farnesyl-diphosphate farnesyltransferase (EC 2.5.1.21) - Nicotiana benthamiana
N;Alternate names: squalene synthetase
C;Species: Nicotiana benthamiana
```

Thu Nov 13 10:01:43 2003

C;Date: 12-Mar-1998 #sequence_revision 12-Mar-1998 #text_change 18-Jun-1999
C;Accession: S71771
R;Hanley, K.M.; Nicolas, O.; Donaldson, T.B.; Smith-Monroy, C.; Robinson, G.W.; Hellmann
Plant Mol. Biol. 30, 1139-1151, 1996
A;Title: Molecular cloning, in vitro expression and characterization of a plant squalene
A;Reference number: S71771; MUID: 96311004; PMID: 8704125
A;Accession: S71771
A;Molecule type: mRNA
A;Residues: 1-411 <HAN>
A;Cross-references: EMBL:U45000; NID: g1184108; PIDN: AAA87048.1; PID: g1184109
A;Note: the authors translated the codon GTC for residue 260 as Ala
C;Superfamily: farnesyl-diphosphate farnesyltransferase
C;Keywords: steroid biosynthesis; transferase

Query Match 78.8%; Score 1617; DB 2; Length 411;
Best Local Similarity 77.5%; Pred. No. 1.6e-118; Mismatches 39; Indels 2; Gaps 1;
Matches 296; Conservative 45

QY 1 MGS LGT MLRY PDDIYPL LKMKRAIEKAEKQIPPEPHMGFCYSMLHKVSRFSFSLVIOQLNT 60
DB 1 MGS LGT MLRY PDDIYPL LKMKRAIEKAEKQIPPEPHMGFCYSMLHKVSRFSFSLVIOQLNT 60
QY 61 ELRNACVYFVLVRLALDTVEDDTSIPTDEKVPILIAFHRHIYDTDMHYSCTGKEYKILMD 120
DB 61 ELRNACVYFVLVRLALDTVEDDTSIPTDEKVPILIAFHRHIYDTDMHYSCTGKEYKILMD 120
QY 61 ELRDVAVCFYLVRLALDTVEDDTSIPTDEKVPILIAFHRHIYDTDMHYSCTGKEYKILMD 120
DB 61 ELRDVAVCFYLVRLALDTVEDDTSIPTDEKVPILIAFHRHIYDTDMHYSCTGKEYKILMD 120
QY 121 QPHHVSAAFLKGYOEAEIEITRRMGAGMAKFCQEVETVDDYDEYCHYVAGLVGL 180
DB 121 QPHHVSAAFLKGYOEAEIEITRRMGAGMAKFCQEVETVDDYDEYCHYVAGLVGL 180
QY 181 SKLFLAAGSEVLTDPWEAISNSMGLFLOKNTIIRDYLEDINEIPKSRMFWPRETWGKYAD 240
DB 181 SKLFLAAGSEVLTDPWEAISNSMGLFLOKNTIIRDYLEDINEIPKSRMFWPRETWGKYAD 240
QY 238 KLEELKYENSTKNSVQCLNEMVTNALMHIEDCLKYMVSLRDPISFRFCAIPQIMAIGTLA 300
DB 238 KLEELKYENSTKNSVQCLNEMVTNALMHIEDCLKYMVSLRDPISFRFCAIPQIMAIGTLA 300
QY 301 LCYNNQVFRGVVLRGLTAKVIDRTKTMADYVGFDFSCMLTKTKVDKNDPNASKT 360
DB 301 LCYNNQVFRGVVLRGLTAKVIDRTKTMADYVGFDFSCMLTKTKVDKNDPNASKT 360
QY 299 MCYDNIEVFRGVVLRGLTAKVIDRTKTMADYVGFDFSCMLTKTKVDKNDPNASKT 358
DB 299 MCYDNIEVFRGVVLRGLTAKVIDRTKTMADYVGFDFSCMLTKTKVDKNDPNASKT 358
QY 361 RLEAVOKLCRDAGVLQNRKSYVNDKQ 387
DB 359 LNRLEAVOKLCRDAGVLQNRKSYVNDKQ 387
DB 358 LNRLEAVOKLCRDAGVLQNRKSYVNDKQ 386

RESULT 8
T05285
farnesyl-diphosphate farnesyltransferase (EC 2.5.1.21) T4L20.230 - Arabidopsis thaliana
N;Alternate names: protein T4L20.230; squalene synthase
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Jul-1999
C;Accession: T05285
R;Bevan, M.; Terry, N.; Ardiles, W.; Buyssehaert, C.; Dasseville, R.; De Clerck, R.; De
ewes, H.W.; Mayer, K.F.X.; Schueller, C.
submitted to the Protein Sequence Database, September 1998
A;Reference number: Z15406
A;Accession: T05285
A;Molecule type: DNA
A;Residues: 1-412 <BEV>
A;Cross-references: EMBL:AL023094
A;Note: T4L20.230
A;Experimental source: cultivar Columbia; BAC clone T4L20
C;Genetics:
A;Map position: 4
A;Introns: 65/3; 109/1; 134/2; 157/3; 207/3; 242/3; 291/3; 317/1; 346/3; 377/2; 391/3
A;Note: T4L20.230
C;Superfamily: farnesyl-diphosphate farnesyltransferase
C;Keywords: transferase

Query Match 78.3%; Score 1605.5; DB 2; Length 412;
Best Local Similarity 77.6%; Pred. No. 1.3e-117; Mismatches 45; Indels 5; Gaps 2;
Matches 302; Conservative 37

QY 1 MGS LGT MLRY PDDIYPL LKMKRAIEKAEKQIPPEPHMGFCYSMLHKVSRFSFSLVIOQLNT 60
DB 1 MGS LGT MLRY PDDIYPL LKMKRAIEKAEKQIPPEPHMGFCYSMLHKVSRFSFSLVIOQLNT 60
QY 61 ELRNACVYFVLVRLA--LDTVEDDTSIPTDEKVPILIAFHRHIYDTDMHYSCTGKEYKIL 118
DB 61 ELRNACVYFVLVRLA--LDTVEDDTSIPTDEKVPILIAFHRHIYDTDMHYSCTGKEYKIL 117
QY 119 MDQPHHVSAAFLKGYOEAEIEITRRMGAGMAKFCQEVETVDDYDEYCHYVAGLVGL 178
DB 118 MDQPHHVSAAFLKGYOEAEIEITRRMGAGMAKFCQEVETVDDYDEYCHYVAGLVGL 177
QY 179 GLSKFLAAGSEVLTDPWEAISNSMGLFLOKNTIIRDYLEDINEIPKSRMFWPRETWGKY 238
DB 179 GLSKFLAAGSEVLTDPWEAISNSMGLFLOKNTIIRDYLEDINEIPKSRMFWPRETWGKY 237
QY 239 ADKLEDKYENSTKNSVQCLNEMVTNALMHIEDCLKYMVSLRDPISFRFCAIPQIMAIGT 298
DB 239 ADKLEDKYENSTKNSVQCLNEMVTNALMHIEDCLKYMVSLRDPISFRFCAIPQIMAIGT 297
QY 238 VKLEDKFKNEEKATRAVQCLNEMVTNALMHIEDCLKSLASLRDPAIFOSCAIPQIVAIGT 297
DB 238 VKLEDKFKNEEKATRAVQCLNEMVTNALMHIEDCLKSLASLRDPAIFOSCAIPQIVAIGT 297
QY 299 LALCYNNQVFRGVVLRGLTAKVIDRTKTMADYVGFDFSCMLTKTKVDKNDPNASKT 358
DB 299 LALCYNNQVFRGVVLRGLTAKVIDRTKTMADYVGFDFSCMLTKTKVDKNDPNASKT 357
QY 359 LNRLEAVOKLCRDAGVLQNRKSYVNDKQ 387
DB 358 LNRLEAVOKLCRDAGVLQNRKSYVNDKQ 386

RESULT 9
T00489
farnesyl-diphosphate farnesyltransferase (EC 2.5.1.21) - maize
N;Alternate names: squalene synthase
C;Species: Zea mays (maize)
C;Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 20-Jun-2000
C;Accession: T00489
R;Hata, S.
submitted to the EMBL Data Library, September 1997
A;Description: CDNA cloning of squalene synthase genes from mono- and dicotyledonous plan
A;Reference number: Z14155
A;Accession: T00489
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-401 <HAT>
A;Cross-references: EMBL:AB007502; PIDN: BAA22558.1
C;Superfamily: farnesyl-diphosphate farnesyltransferase
C;Keywords: transferase

Query Match 70.5%; Score 1446; DB 2; Length 401;
Best Local Similarity 70.4%; Pred. No. 3.5e-105; Mismatches 49; Indels 6; Gaps 2;
Matches 266; Conservative 57

QY 1 MGS LGT MLRY PDDIYPL LKMKRAIEKAEKQIPPEPHMGFCYSMLHKVSRFSFSLVIOQLNT 60
DB 1 MGS LGT MLRY PDDIYPL LKMKRAIEKAEKQIPPEPHMGFCYSMLHKVSRFSFSLVIOQLNT 56
QY 61 ELRNACVYFVLVRLALDTVEDDTSIPTDEKVPILIAFHRHIYDTDMHYSCTGKEYKILMD 120
DB 61 ELRNACVYFVLVRLALDTVEDDTSIPTDEKVPILIAFHRHIYDTDMHYSCTGKEYKILMD 116
QY 121 QPHHVSAAFLKGYOEAEIEITRRMGAGMAKFCQEVETVDDYDEYCHYVAGLVGL 180
DB 121 QPHHVSAAFLKGYOEAEIEITRRMGAGMAKFCQEVETVDDYDEYCHYVAGLVGL 176
QY 177 KFRHVSTAFLELGGYQKAEIEVTRMGAGMAKFCQEVETVDDYDEYCHYVAGLVGL 240
DB 177 KFRHVSTAFLELGGYQKAEIEVTRMGAGMAKFCQEVETVDDYDEYCHYVAGLVGL 234
QY 241 KLEDKYENSTKNSVQCLNEMVTNALMHIEDCLKYMVSLRDPISFRFCAIPQIMAIGTLA 300
DB 235 KLEDKYENSTKNSVQCLNEMVTNALMHIEDCLKYMVSLRDPISFRFCAIPQIMAIGTLA 294

Qy 301 LCYNNEQVFRGVKLRGLTAKVIDRTKTMADYVGAIFYDFSCMLKTKVDKNDPNASKTLN 360
Db 295 LCYNVHVFGRGVKLRGLTAKVIDRTKTMADYVGAIFYDFSCMLKTKVDKNDPNASKTLN 354
Qy 361 RLEAVQKLCRDAGVQLNQR 378
Db 355 RVDSIKOTCKSSGLVKQR 372

RESULT 10
T02167
farnesyl-diphosphate farnesyltransferase (EC 2.5.1.21) - common tobacco
N;Alternate names: squalene synthase
C;Species: Nicotiana tabacum (common tobacco)
C;Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 23-Jul-1999
C;Accession: T02167
R;Devarene, T.P.; Shin, D.H.; Back, K.; Yin, S.; Chappell, J.
submitted to the EMBL Data Library, May 1996
A;Description: Genomic tobacco squalene synthase gene sequence.
A;Reference number: 214605
A;Accession: T02167
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-376 <DEV>
A;Cross-references: EMBL:U59683; NID:gl389892; PIDN:AAB02945.1; PID:gl389893
C;Genetics:
A;Gene: TSS
A;Introns: 65/3; 79/3; 110/3; 135/3; 158/3; 206/3; 255/3; 281/3; 310/3; 341/3; 355/3
C;Superfamily: farnesyl-diphosphate farnesyltransferase
C;Keywords: steroid biosynthesis; transferase

Query Match 70.1%; Score 1437.5; DB 2; Length 376;
Best Local Similarity 70.4%; Pred. No. 1.5e-104;
Matches 269; Conservative 42; Mismatches 34; Indels 37; Gaps 2;

Qy 1 MGSIGTMLRYPDDIYPLLLKWKRAIEKAQIPEPHWGFCSYMLHKVSRFSFSLVIQQLNT 60
Db 1 MGSURAILKNPDDIYPLVKKLAARHAEKQIPPSPHWGFCSYMLHKVSRFSFSLVIQQLPV 60

Qy 61 ELRNVAVFYLVRALDVTVEDDTSIPTDEKVPILIAFHRHIYDTHWYSCGTKEYKILMD 120
Db 61 ELRDAVCFILVRALDVTVEDDTSIPTDKVPILISFQHVYDREWHFSCGTKEYKVLMD 120

Qy 121 QFHHVSAFLELEKGYQEAIEITRRMGAGMAKFCIQEVTVDYDEYCHYVAGLVGLGL 180
Db 121 QFHHVSTAFLELRKHYYQQAIEDITRRMGAGMAKFCIQEVTVDYDEYCHYVAGLVGLGL 180

Qy 181 SKLFLAAGSEVLTDPWEAISNMGFLQKTNIIIRDYLEINEIPKSRMFHPREIWKYAD 240
Db 181 SKLFLASGKEDLASD--SLNSMGLFLQ----- 206

Qy 241 KLEDLKYEENTNKSVOCLNEMVTNALMHIEDCLKYMVSLRDPSPFRCAIPQINAIQTGLA 300
Db 207 ---ELKTEDNSAKAVQCLNDMVTNALSHVEDCLTYNGALRDPSPFRCAIPQINAIQTGLA 263

Qy 301 LCYNNEQVFRGVKLRGLTAKVIDRTKTMADYVGAIFYDFSCMLKTKVDKNDPNASKTLN 360
Db 264 MCDNIEVFRGVKLRGLTAKVIDQRTIADYVGAFFDFSCMLKSKVNNNDPNATKTLK 323

Qy 361 RLEAVQKLCRDAGVQLNQRKSYV 382
Db 324 RLEAILKTCRDSGTLNRKRSYI 345

RESULT 11
S52075
farnesyl-diphosphate farnesyltransferase (EC 2.5.1.21) - mouse
N;Alternate names: squalene synthase
C;Species: Mus musculus (house mouse)
C;Date: 14-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 20-Jun-2000
C;Accession: S52075
R;Inoue, T.; Oeumi, T.; Hata, S.
Biochim. Biophys. Acta 1260, 49-54, 1995

A;Title: Molecular cloning and functional expression of a cDNA for mouse squalene synthase
A;Reference number: S52075; MUID:95032793; PMID:7999794
A;Accession: S52075
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-416 <INO>
A;Cross-references: EMBL:D29016; NID:g468456; PIDN:BAA06102.1; PID:g511932
C;Superfamily: farnesyl-diphosphate farnesyltransferase
C;Keywords: transferase

Query Match 41.3%; Score 848; DB 2; Length 416;
Best Local Similarity 46.9%; Pred. No. 1.7e-58;
Matches 168; Conservative 70; Mismatches 108; Indels 12; Gaps 5;

Qy 8 LRYPDIDYPLLLKWKRA-----IEKAQIPEPHWGFCSYMLHKVSRFSFSLVIQQLNTEL 62
Db 7 LGHPEEFYLLRFRMGGRNRFIPKMD-QDSLSLTKCYKYLQNTSFAVIALQDGI 65

Qy 63 RNACVVFYLVRALDVTVEDDTSIPTDEKVPILIAFHRHIYDTHWYSCGTKEYKILMDQF 122
Db 66 RHAICVFYLVRALDVTVEDDMSISVEKKIPLLCNFHTFLYDPEWRFTESEKDRQVLEDF 125

Qy 123 HHVSAAFLELEKGYQEAIEITRRMGAGMAKFCIQEVTVDYDEYCHYVAGLVGLGLSK 182
Db 126 PTISLSEFRNLAEKYQTVIDDICHOMGCGMAEFVDKDVTSKQDMWKYCHYVAGLVGLSLR 185

Qy 183 LFLAAGSE--VLTPDWEAISNMGFLQKTNIIIRDYLEINEIPKSRMFHPREIWKYAD 240
Db 186 LFSASEPFDIVGEDIEC-ANSMGLFLQKTNIIIRDYLEDOQE---GRKFPQEWGVIYK 241

Qy 241 KLEDLKYEENTNKSVOCLNEMVTNALMHIEDCLKYMVSLRDPSPFRCAIPQINAIQTGLA 300
Db 242 KLEDFAKPENVDVAVQCLNELITNTLQHIPDLVLTYSRLRNSQSVFNCAIPQVMAIATLA 301

Qy 301 LCYNNEQVFRGVKLRGLTAKVIDRTKTMADYVGAIFYDFSCMLKTKVDKNDPNASKT 358
Db 302 ACYNNQOVFGVVKIRKQAVTLLMDATNPAVKAIYIEEYHRIPIPSDPSSSKT 359

RESULT 12
A45105
farnesyl-diphosphate farnesyltransferase (EC 2.5.1.21) - rat
N;Alternate names: squalene synthase
C;Species: Rattus norvegicus (Norway rat)
C;Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 18-Jun-1999
C;Accession: A45105
R;McKenzie, T.L.; Jiang, G.; Straubhaar, J.R.; Conrad, D.G.; Shechter, I.
J. Biol. Chem. 267, 21368-21374, 1992
A;Title: Molecular cloning, expression, and characterization of the cDNA for the rat hep
A;Reference number: A45105; MUID:93016074; PMID:1400448
A;Accession: A45105
A;Status: preliminary
A;Molecule type: mRNA; protein
A;Residues: 1-416 <MCK>
A;Cross-references: GB:M95591; NID:g207070; PIDN:AAA42179.1; PID:g207071
A;Experimental source: hepatoma cell line H35
A;Note: sequence extracted from NCBI backbone (NCBIN:116694, NCBI:P:116696)
C;Superfamily: farnesyl-diphosphate farnesyltransferase
C;Keywords: transferase; transmembrane protein

Query Match 41.2%; Score 845; DB 2; Length 416;
Best Local Similarity 46.5%; Pred. No. 3e-58;
Matches 166; Conservative 71; Mismatches 108; Indels 12; Gaps 5;

Qy 8 LRYPDIDYPLLLKWKRA-----IEKAQIPEPHWGFCSYMLHKVSRFSFSLVIQQLNTEL 62
Db 7 LGHPEEFYLLRFRMGGRNRFIPKMDR-SLSLSLTKCYKYLQNTSFAVIALQDGI 65

Qy 63 RNACVVFYLVRALDVTVEDDTSIPTDEKVPILIAFHRHIYDTHWYSCGTKEYKILMDQF 122
Db 66 RHAICVFYLVRALDVTVEDDMSISVEKKIPLLCNFHTFLYDPEWRFTESEKDRQVLEDF 125

Qy 123 HHVSAAFLELEKGYQEAIEITRRMGAGMAKFCIQEVTVDYDEYCHYVAGLVGLGLSK 182

```

Db      126 PTISLEFNLAEKQTVIADICHRMGCGMAEFLNKQDWDKICYHYVAGLVGIGLSR 185
Qy      183 LFLAAGSE--VLTDPWEAISNMGFLQKTNIIIRDYLEINEIPKSRMFWPREINGKYAD 240
Db      186 LFSASEFEDPIVGEDTEC-ANSMGLFLQKTNIIIRDYLEDOQ---GQFQFQVWGVKYYK 241
Qy      241 KLEDLKYEENTNKSVOCLNEMVTNALMHIEDCLKYMVSLRDPISIFRCAIPQIMAIQTAL 300
Db      242 KLEDVFKPENVDVAVKCLNELITNALHHIPDVTYLSRLRNSQSVFNFCALPQVMAIATLA 301
Qy      301 LCYNNEQVFRGVVKLRRLGTAKVIDRTKTMADVGAFYDFSCMLKTKVKDKNPNASK 357
Db      302 ACYNHNVQVKGKVKIRKQAVTLMMDATNMPAVKAIYQYIEEYIYHRVFNPSDPSASK 358

RESULT 13
A:Species: Homo sapiens (man)
N:Alternate names: squalene synthase
C:Species: Homo sapiens (man)
C>Date: 03-Mar-1994 #sequence_revision 18-Nov-1994 #text_change 18-Jun-1999
C:Accession: A45998
R:Jiang, G.; McKenzie, T.L.; Conrad, D.G.; Shechter, I.
J. Biol. Chem. 268, 12818-12824, 1993
A:Title: Transcriptional regulation by lovastatin and 25-hydroxycholesterol in HepG2 cell
A:Reference number: A45998; MUID:93286128; PMID:7685352
A:Accession: A45998
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-417 <L1A>
A:Cross-references: GB:L06105; NID:g307431; PIDN:AAA36645.1; PID:g307432
A:Experimental source: hepatoma cell line HepG2
A>Note: sequence extracted from NCBI backbone (NCBIN:133625, NCBIIP:133626)
R:Robinson, G.W.; Tsay, Y.H.; Kienzie, B.K.; Smith-Monroy, C.A.; Bishop, R.W.
Mol. Cell. Biol. 13, 2706-2717, 1993
A:Title: Conservation between human and fungal squalene synthetases: similarities in str
A:Reference number: A48057; MUID:93233634; PMID:8474436
A:Accession: A48057
A>Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-417 <ROB>
A:Cross-references: GB:L06070; NID:g292509; PIDN:AAA60582.1; PID:g292510
A>Note: sequence extracted from NCBI backbone (NCBIN:129790, NCBIIP:129791)
C:Superfamily: farnesyl-diphosphate farnesyltransferase
C:Keywords: transferase; transmembrane protein

Query Match      40.0%; Score 820; DB 2; Length 417;
Best Local Similarity 45.2%; Pred. No. 2.7e-56;
Matches 165; Conservative 69; Mismatches 121; Indels 10; Gaps 4;

Qy      8 LRYPDIDYPLLMK-----KRAIEKAEKQIPEPHWGCYSLMLHKVSRFSLSLVIQQLNTELR 63
Db      7 LGHPEEFYNLVRPRIGGKRWPKMDQDLSLSSSLKTCYKYLNTQTSRFAAVIQALDGENR 66
Qy      64 NAVCVFLVLRALDVTVEDDTSIPTDEKVPILIAFHRIHYDTDWHYSCGTKEYKILMDQPH 123
Db      67 NAVCIFYLVRALDVTLEDMMTISVEKKVPLLNHFHSLFYQPDWRFWESKEKDRQVLEDP 126
Qy      124 HVSAAFLELEKGYQEAIEETRRMGAGMAKFIQCEYVETVDDYDEYCHYVAGLVGLSKL 183
Db      127 TISLEFNLAEKQTVIADICRMGIGMAEFLDKHVTSEQWDKYCHYVAGLVGIGLSRL 186
Qy      184 FLAAGSE--VLTDPWEAISNMGFLQKTNIIIRDYLEINEIPKSRMFWPREINGKYADK 241
Db      187 FSASEFEDPIVGEDTER-ANSMGLFLQKTNIIIRDYLEDOQ---GGREFPQVWVSRVVK 242
Qy      242 LEDLKYEENTNKSVOCLNEMVTNALMHIEDCLKYMVSLRDPISIFRCAIPQIMAIQTAL 301
Db      243 LGDPAKENIDLAQCCLNELITNALHHIPDVTYLSRLRNSQSVFNFCALPQVMAIATLA 302
Qy      302 CYNNEQVFRGVVKLRRLGTAKVIDRTKTMADVGAFYDFSCMLKTKVKDKNPNASKTLNR 361
Db      303 CYNNOQVFKGAVKIRKQAVTLMMDATNMPAVKAIYQYIEEYIYHRIPDSDPSSSKTRQI 362
Qy      362 LEAVQ 366
Db      363 ISTIR 367

RESULT 15
squalene synthase - human
C:Species: Homo sapiens (man)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 18-Jun-1999
C:Accession: I52090
R:Soltis, D.A.; McMahon, G.; Caplan, S.L.; Dudas, D.A.; Chamberlin, H.A.; Vattay, A.; Doi
Arch. Biochem. Biophys. 316, 713-723, 1995
A:Title: Expression, purification, and characterization of the human squalene synthase:
A:Reference number: I52090; MUID:95168856; PMID:7864626
A:Accession: I52090
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-417 <RES>

```

```

Db      303 CYNNOQVFKGAVKIRKQAVTLMMDATNMPAVKAIYQYIEEYIYHRIPDSDPSSSKTRQI 362
Qy      362 LEAVQ 366
Db      363 ISTIR 367

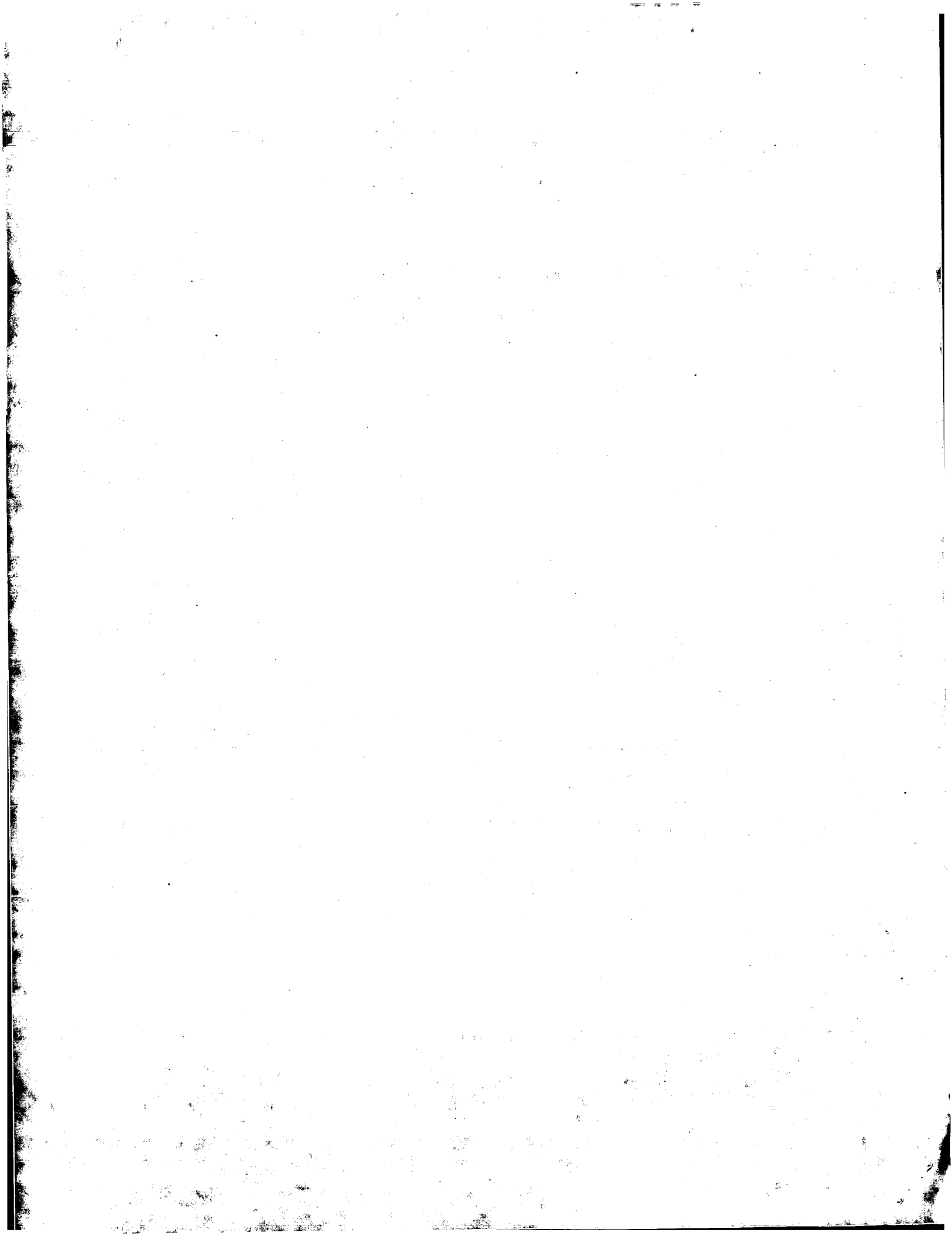
RESULT 14
farnesyl-diphosphate farnesyltransferase (EC 2.5.1.21), hepatic - human
N:Alternate names: squalene synthase
C:Species: Homo sapiens (man)
C>Date: 15-Mar-1996 #sequence_revision 15-Mar-1996 #text_change 18-Jun-1999
C:Accession: I38245
R:Summers, C.; Karst, F.; Charles, A.D.
Gene 136, 185-192, 1993
A:Title: Cloning, expression and characterisation of the cDNA encoding human hepatic squ
A:Reference number: I38245; MUID:94123996; PMID:8294001
A:Accession: I38245
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-417 <RES>
A:Cross-references: EMBL:X69141; NID:g435676; PIDN:CAA48896.1; PID:g435677
C:Superfamily: farnesyl-diphosphate farnesyltransferase
C:Keywords: liver; membrane protein; steroid biosynthesis; transferase

Query Match      40.0%; Score 820; DB 2; Length 417;
Best Local Similarity 45.2%; Pred. No. 2.7e-56;
Matches 165; Conservative 69; Mismatches 121; Indels 10; Gaps 4;

Qy      8 LRYPDIDYPLLMK-----KRAIEKAEKQIPEPHWGCYSLMLHKVSRFSLSLVIQQLNTELR 63
Db      7 LGHPEEFYNLVRPRIGGKRWPKMDQDLSLSSSLKTCYKYLNTQTSRFAAVIQALDGENR 66
Qy      64 NAVCVFLVLRALDVTVEDDTSIPTDEKVPILIAFHRIHYDTDWHYSCGTKEYKILMDQPH 123
Db      67 NAVCIFYLVRALDVTLEDMMTISVEKKVPLLNHFHSLFYQPDWRFWESKEKDRQVLEDP 126
Qy      124 HVSAAFLELEKGYQEAIEETRRMGAGMAKFIQCEYVETVDDYDEYCHYVAGLVGLSKL 183
Db      127 TISLEFNLAEKQTVIADICRMGIGMAEFLDKHVTSEQWDKYCHYVAGLVGIGLSRL 186
Qy      184 FLAAGSE--VLTDPWEAISNMGFLQKTNIIIRDYLEINEIPKSRMFWPREINGKYADK 241
Db      187 FSASEFEDPIVGEDTER-ANSMGLFLQKTNIIIRDYLEDOQ---GGREFPQVWVSRVVK 242
Qy      242 LEDLKYEENTNKSVOCLNEMVTNALMHIEDCLKYMVSLRDPISIFRCAIPQIMAIQTAL 301
Db      243 LGDPAKENIDLAQCCLNELITNALHHIPDVTYLSRLRNSQSVFNFCALPQVMAIATLA 302
Qy      302 CYNNEQVFRGVVKLRRLGTAKVIDRTKTMADVGAFYDFSCMLKTKVKDKNPNASKTLNR 361
Db      303 CYNNOQVFKGAVKIRKQAVTLMMDATNMPAVKAIYQYIEEYIYHRIPDSDPSSSKTRQI 362
Qy      362 LEAVQ 366
Db      363 ISTIR 367

RESULT 15
squalene synthase - human
C:Species: Homo sapiens (man)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 18-Jun-1999
C:Accession: I52090
R:Soltis, D.A.; McMahon, G.; Caplan, S.L.; Dudas, D.A.; Chamberlin, H.A.; Vattay, A.; Doi
Arch. Biochem. Biophys. 316, 713-723, 1995
A:Title: Expression, purification, and characterization of the human squalene synthase:
A:Reference number: I52090; MUID:95168856; PMID:7864626
A:Accession: I52090
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-417 <RES>

```

GenCore version S.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 12, 2003, 16:20:45 ; Search time 17 Seconds
(without alignments)
1073.315 Million cell updates/sec

Title: US-10-024-130A-6

Perfect score: 2051

Sequence: 1 MGSLGTMRLYPDDIYPLLM.....CRDAGVLQNRKSYVNDKGQP 388

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | ID | Description |
|------------|-------|---------------|--------|----|-------------|
| 1 | 2051 | 100.0 | 410 | 1 | FDFT_ARATH |
| 2 | 1617 | 78.8 | 411 | 1 | FDFT_NICBE |
| 3 | 848 | 41.3 | 416 | 1 | FDFT_MOUSE |
| 4 | 845 | 41.2 | 416 | 1 | FDFT_RAT |
| 5 | 820 | 40.0 | 417 | 1 | FDFT_HUMAN |
| 6 | 747 | 36.4 | 448 | 1 | FDFT_CANAL |
| 7 | 721 | 35.2 | 444 | 1 | FDFT_YEAST |
| 8 | 709.5 | 34.6 | 443 | 1 | FDFT_CANGA |
| 9 | 700.5 | 34.2 | 460 | 1 | FDFT_SCHPO |
| 10 | 617.5 | 30.1 | 572 | 1 | FDFT_USTMA |
| 11 | 122.5 | 6.0 | 302 | 1 | CRTB_MYCTU |
| 12 | 118 | 5.8 | 410 | 1 | PSY_MAIZE |
| 13 | 117 | 5.7 | 398 | 1 | PSY_DAUCA |
| 14 | 115 | 5.6 | 309 | 1 | CRTB_PANAN |
| 15 | 112.5 | 5.5 | 279 | 1 | Y4AC_RHISN |
| 16 | 109.5 | 5.3 | 339 | 1 | CRTB_RHOCA |
| 17 | 109.5 | 5.3 | 674 | 1 | YB66_YEAST |
| 18 | 107.5 | 5.2 | 309 | 1 | CRTB_ERWHE |
| 19 | 107 | 5.2 | 337 | 1 | CRTB_SYNY3 |
| 20 | 106.5 | 5.2 | 342 | 1 | CRTB_STRGR |
| 21 | 105.5 | 5.1 | 289 | 1 | CRTB_THETH |
| 22 | 104.5 | 5.1 | 292 | 1 | Y4AD_RHISN |
| 23 | 104 | 5.1 | 423 | 1 | PSY_NARPS |
| 24 | 103 | 5.0 | 583 | 1 | GLS2_CABEL |
| 25 | 102.5 | 5.0 | 412 | 1 | PSY2_LYCES |
| 26 | 102.5 | 5.0 | 422 | 1 | PSY_ARATH |
| 27 | 102 | 5.0 | 309 | 1 | CRTY_SPIPL |
| 28 | 101.5 | 4.9 | 310 | 1 | PSY2_LYCES |
| 29 | 100.5 | 4.9 | 603 | 1 | RAEP_YEAST |
| 30 | 99.5 | 4.9 | 795 | 1 | ENPL_CHICK |
| 31 | 99 | 4.8 | 419 | 1 | PSY_CAPAN |
| 32 | 97.5 | 4.8 | 301 | 1 | CRTB_AGRAU |
| 33 | 97 | 4.7 | 722 | 1 | GREX_CHLPN |

| | | | | | | |
|----|------|-----|------|---|------------|--------------------|
| 34 | 96.5 | 4.7 | 803 | 1 | ENPL_HUMAN | P14625 homo sapien |
| 35 | 96.5 | 4.7 | 804 | 1 | ENPL_CANFA | P41148 canis famil |
| 36 | 96.5 | 4.7 | 804 | 1 | ENPL_PIG | Q29092 sus scrifa |
| 37 | 96.5 | 4.7 | 1110 | 1 | YNO3_YEAST | P53893 saccharomyc |
| 38 | 95.5 | 4.7 | 459 | 1 | YA03_HUMAN | O60809 homo sapien |
| 39 | 95 | 4.6 | 691 | 1 | YIO4_YEAST | P40460 saccharomyc |
| 40 | 95 | 4.6 | 1272 | 1 | Y228_METJA | Q60287 methanococc |
| 41 | 95 | 4.6 | 1661 | 1 | YBT1_YEAST | P32386 saccharomyc |
| 42 | 94.5 | 4.6 | 422 | 1 | PSY_CUCME | P49293 cucumis mel |
| 43 | 94.5 | 4.6 | 1003 | 1 | TP3A_MOUSE | O70157 mus musculu |
| 44 | 94 | 4.6 | 917 | 1 | MSH2_DROME | P43248 drosophila |
| 45 | 94 | 4.6 | 1005 | 1 | RASO_METJA | Q58718 methanococc |

ALIGNMENTS

RESULT 1
FDFT_ARATH STANDARD; PRT; 410 AA.
AC P53799;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Farnesyl-diphosphate farnesyltransferase (EC 2.5.1.21) (Squalene synthetase) (SQS) (SS) (FPP:FPP farnesyltransferase).
GN SQS1 OR AT4G34640 OR T4L20.220.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=95199342; PubMed=7892265;
RA Nakashima T., Inoue T., Oka A., Nishino T., Osumi T., Hata S.;
RT "Cloning, expression, and characterization of cDNAs encoding Arabidopsis thaliana squalene synthase."
RL Proc. Natl. Acad. Sci. U.S.A. 92:2328-2332(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Landsberg erecta;
RA Kribil R., Boronat A., Karst F.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Kribil R., Arro M., del Arco A., Gonzalez V., Balcells L.L.,
RA Delorme D., Ferrer A., Karst F., Boronat A.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Connolly E.L., Learned R.M.;
RT "Isolation and characterization of squalene synthase from Arabidopsis thaliana."
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083488; PubMed=10617198;
RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
RA Pohl T., Dueterhoeft A., Stiekema W., Entian K.-D., Terry N.,
RA Harris B., Anson W., Brandt P., Griwell L.A., Rieger M.,
RA Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,
RA Kreis M., Delsen J., Puigdomenech P., Watson M., Schmidheini T.,
RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA Vos P., Hohseil J., Zimmermann W., Wedler H., Ridley P.,
RA Langham S.-A., McCullagh B., Wilham L., Robben J.,
RA Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,
RA Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
RA Weitzensgger T., Bothe G., Rampsperger U., Hilbert H., Braun M.,
RA Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,
RA Mooijman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,

| | | | | | | | | | | | |
|-------------|---|---------|------|--------------|----|------------|----|--------|----|------|----|
| | Berneiser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H., De Keyser A., Buyschaert C., Gielen J., Villarroel R., De Clercq R., Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S., Clark L., Doggett J., Hall S., Kay M., Lennard N., McVay K., Mayes R., Pettett A., Rajandream M.A., Lyne M., Benes V., Reckmann S., Borkova D., Blocker H., Scharfe M., Grimm M., Loehner T.-H., Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Fartmann B., Grandrath K., Dauner D., Herzl A., Neumann S., Argirou A., Vitale D., Ligouri R., Piravandi E., Massenot O., Quigley F., Clabaud G., Muendlein A., Feilber R., Schnabl S., Hiller J., Schmidt W., Lecharny A., Aubourg S., Chedof F., Cooke R., Berger C., Monfort A., Casacuberta E., Gibbons T., Weber N., Vandenbol M., Bagues M., Terol J., Torres A., Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T., Reijnders L., Schwarz S., Scholler P., Heber S., Frances P., Bieleke C., Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S., Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K., Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L., Saknon M., Murray J., Sheet P., Cordes M., Abu-Threideh J., Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J., Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D., Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K., Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L., Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D., Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshua C., Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C., Ma P., Zhong J., Preston R., Vill D., Shekher M., Matero A., Shah R., Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S., Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A., Chen E., Marra M., Martienssen R., McCombie W.R.; "Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana." Nature 402:769-777(1999). | Matches | 388; | Conservative | 0; | Mismatches | 0; | Indels | 0; | Gaps | 0; |
| QY | 1 MGSGLTMRYPDDIYPLLLKMKRAIEKAQIPEPHMGFCYSMLHKVSRFSFSLVIQOLNT 60 | | | | | | | | | | |
| DB | 1 MGSGLTMRYPDDIYPLLLKMKRAIEKAQIPEPHMGFCYSMLHKVSRFSFSLVIQOLNT 60 | | | | | | | | | | |
| QY | 61 ELRNVACVFLVLRALDVTVEDDTSIPTDEKVPILIAFHRIHYDTDWHYSCGTKEYKILMD 120 | | | | | | | | | | |
| DB | 61 ELRNVACVFLVLRALDVTVEDDTSIPTDEKVPILIAFHRIHYDTDWHYSCGTKEYKILMD 120 | | | | | | | | | | |
| QY | 121 QFHVSAFLELEKGYOEATEETTRMGAGMAKFIQOEVTVDYDYBYCHYVAGLVGLGL 180 | | | | | | | | | | |
| DB | 121 QFHVSAFLELEKGYOEATEETTRMGAGMAKFIQOEVTVDYDYBYCHYVAGLVGLGL 180 | | | | | | | | | | |
| QY | 181 SKFLAAGSEVLTPDWEAISNSMGLFLQKTNIIIRDYLEDINETPKSRMFWPREIWGYAD 240 | | | | | | | | | | |
| DB | 181 SKFLAAGSEVLTPDWEAISNSMGLFLQKTNIIIRDYLEDINETPKSRMFWPREIWGYAD 240 | | | | | | | | | | |
| QY | 241 KLEDIKYEENTNKSQVCLNEMVTNALMHIEDCLKYVSLRDPSTFRFCALPQIMAICTLA 300 | | | | | | | | | | |
| DB | 241 KLEDIKYEENTNKSQVCLNEMVTNALMHIEDCLKYVSLRDPSTFRFCALPQIMAICTLA 300 | | | | | | | | | | |
| QY | 301 LCYNNEQVFRGWKLRGLTAKVIDRTKTADVYGAFYDFSCMLTKTKVDKNDPNASKTLN 360 | | | | | | | | | | |
| DB | 301 LCYNNEQVFRGWKLRGLTAKVIDRTKTADVYGAFYDFSCMLTKTKVDKNDPNASKTLN 360 | | | | | | | | | | |
| QY | 361 RLEAVOKLCRDAGVLQNRKSYVNDKGQP 388 | | | | | | | | | | |
| DB | 361 RLEAVOKLCRDAGVLQNRKSYVNDKGQP 388 | | | | | | | | | | |
| RESULT 2 | | | | | | | | | | | |
| FDPRT_NICBE | STANDARD; PRT; 411 AA. | | | | | | | | | | |
| ID | FDPRT_NICBE | | | | | | | | | | |
| AC | P53800; | | | | | | | | | | |
| DT | 01-OCT-1996 (Rel. 34, Created) | | | | | | | | | | |
| DT | 01-OCT-1996 (Rel. 34, Last sequence update) | | | | | | | | | | |
| DT | 28-FEB-2003 (Rel. 41, Last annotation update) | | | | | | | | | | |
| DE | Farnesyl-diphosphate farnesyltransferase (EC 2.5.1.21) (Squalene synthetase) (SQS) (SS) (FPP:Fpp farnesyltransferase). | | | | | | | | | | |
| OS | Nicotiana benthamiana. | | | | | | | | | | |
| OC | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; lamids; Solanales; Solanaceae; Nicotiana. | | | | | | | | | | |
| OX | NCBI_TaxID=4100; | | | | | | | | | | |
| RN | [1] | | | | | | | | | | |
| RP | SEQUENCE FROM N.A. | | | | | | | | | | |
| RX | MEDLINE=96311004; PubMed=8704125; | | | | | | | | | | |
| RA | Hanley K.M., Nicolas O., Donaldson T.B., Smith-Monroy C., Robinson G.W., Hellmann G.M.; | | | | | | | | | | |
| RT | "Molecular cloning, in vitro expression and characterization of a plant squalene synthetase cDNA."; | | | | | | | | | | |
| RL | Plant Mol. Biol. 30:1139-1151(1996). | | | | | | | | | | |
| CC | -!- CATALYTIC ACTIVITY: 2 farnesyl diphosphate = diphosphate + presqualene diphosphate. | | | | | | | | | | |
| CC | -!- CATALYTIC ACTIVITY: Presqualene diphosphate + NADPH = diphosphate + squalene + NADP(+). | | | | | | | | | | |
| CC | -!- COFACTOR: MAGNESIUM (BY SIMILARITY). | | | | | | | | | | |
| CC | -!- PATHWAY: Critical branch point enzyme of isoprenoid and cholesterol biosynthesis. | | | | | | | | | | |
| CC | -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic reticulum (By similarity). | | | | | | | | | | |
| CC | -!- SIMILARITY: BELONGS TO THE PHYTOENE/SQUALENE SYNTHETASE FAMILY. | | | | | | | | | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch). | | | | | | | | | | |
| CC | EMBL; U46000; AAA87048.1; - | | | | | | | | | | |
| DR | EMBL; U46000; AAA87048.1; - | | | | | | | | | | |

| | | | |
|----------|--|--|--------------|
| QY | 1 | MSLGTMLRYDDIYPLLLKMKRAIEKAQKQIPPEPHWGFCYSMLHKVSRFSFLVIQQLNT | 60 |
| DB | 1 | MSLGTMLRYDDIYPLLLKMKRAIEKAQKQIPPEPHWGFCYSMLHKVSRFSFLVIQQLNT | 60 |
| QY | 61 | ELRNAVCVFLVLRALDVEDDTSIPTDEKVPILIAFHRHIYDTDMHYSCGTKEYKILMD | 120 |
| DB | 61 | ELRNAVCVFLVLRALDVEDDTSIPTDEKVPILIAFHRHIYDTDMHYSCGTKEYKILMD | 120 |
| QY | 121 | QPHVSAAPLEKGYQAIIEITRMGAGMAKFCQEVETVDDYDEYCHVAGLVGL | 180 |
| DB | 121 | QPHVSAAPLEKGYQAIIEITRMGAGMAKFCQEVETVDDYDEYCHVAGLVGL | 180 |
| QY | 181 | SKLFLAAGSEVLTPDWEAISNMGFLQKTNIIIRDYLEIDINEIPKSRMFWPREIWGYAD | 240 |
| DB | 181 | SKLFLAAGSEVLTPDWEAISNMGFLQKTNIIIRDYLEIDINEIPKSRMFWPREIWGYAD | 240 |
| QY | 241 | KLEDLYKYEENTNKSVOCLNEMVTNLMHIEDCLKYNVSLRDPSPFRFCAIPQIMAIGTLA | 300 |
| DB | 241 | KLEDLYKYEENTNKSVOCLNEMVTNLMHIEDCLKYNVSLRDPSPFRFCAIPQIMAIGTLA | 300 |
| QY | 301 | LCYNNEQVFRGVVYKLRGLTAKVIDRTKTMADYVGAFFDFSCMLTKYVDKNDPNASKTLN | 360 |
| DB | 301 | LCYNNEQVFRGVVYKLRGLTAKVIDRTKTMADYVGAFFDFSCMLTKYVDKNDPNASKTLN | 360 |
| QY | 361 | RLNAVQKLCRDAGVLQNRKSYVNDKGP | 388 |
| DB | 361 | RLNAVQKLCRDAGVLQNRKSYVNDKGP | 388 |
| RESULT 2 | | | |
| FDFT | NICBCE | STANDARD; | PRT; 411 AA. |
| ID | FDFT | NICBCE | STANDARD; |
| AC | P53800; | | |
| DT | 01-OCT-1996 (Rel. 34, Created) | | |
| DT | 01-OCT-1996 (Rel. 34, Last sequence update) | | |
| DT | 28-FEB-2003 (Rel. 41, Last annotation update) | | |
| DE | Farnesyl-diphosphate farnesyltransferase (EC 2.5.1.21) (Squalene synthetase) (SQS) (SS) (FPP:FPP farnesyltransferase). | | |
| OS | Nicotiana benthamiana. | | |
| OC | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; | | |
| OC | Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; | | |
| OC | Asteridae; lamids; Solanales; Solanaceae; Nicotiana. | | |
| OX | NCBI_TaxID=4100; | | |
| RN | [1] | | |
| RP | SEQUENCE FROM N.A. | | |
| RP | MEDLINE=96311004; PubMed=8704125; | | |
| RA | Hanley K.M., Nicolas O., Donaldson T.B., Smith-Monroy C., | | |
| RA | Robinson G.W., Hellmann G.M.; | | |
| RT | "Molecular cloning, in vitro expression and characterization of a plant squalene synthetase cDNA." | | |
| RL | Plant Mol. Biol. 30:1139-1151(1996). | | |
| CC | -!- CATALYTIC ACTIVITY: 2 farnesyl diphosphate = diphosphate + presqualene diphosphate. | | |
| CC | -!- CATALYTIC ACTIVITY: Presqualene diphosphate + NADPH = diphosphate + squalene + NADP(+). | | |
| CC | -!- COFACTOR: MAGNESIUM (By similarity). | | |
| CC | -!- PATHWAY: Critical branch point enzyme of isoprenoid and cholesterol biosynthesis. | | |
| CC | -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic reticulum (By similarity). | | |
| CC | -!- SIMILARITY: BELONGS TO THE PHYTOENE/SQUALENE SYNTHETASE FAMILY. | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch). | | |
| CC | EMBL; U46000; AAA87048.1; - | | |


```

DR PIR; S71771; S71771.
DR InterPro; IPR002060; Squ/phyt synthase.
DR Pfam; PF00494; SQS_PSY; 1.
DR TIGRFAMs; TIGR01559; squa1 synth; 1.
DR PROSITE; PS01044; SQUALEN_PHYTOEN_SYN_1; 1.
DR PROSITE; PS01045; SQUALEN_PHYTOEN_SYN_2; 1.
KW Multifunctional enzyme; Transferrase; Oxidoreductase; NADP; Magnesium;
KW Isoprene biosynthesis; Transmembrane; Endoplasmic reticulum.
FT TRANSMEM 281 301
FT TRANSMEM 388 408
SQ SEQUENCE 411 AA; 47082 MW; 5DE209E9BB70A42A CRC64;

Query Match      78.8%; Score 1617; DB 1; Length 411;
Best Local Similarity 77.5%; Pred. No. 2.2e-117;
Matches 296; Conservative 45; Mismatches 39; Indels 2; Gaps 1;

QY 1 MGSIGTMRVDDPIYPLKMKRAEKAQIPPEPHMGFCYSMLHKVSRFSLSVIQOINT 60
DB 1 MGSRLAIKPNPDYPLVPLKLAARHAEKQIPPSNMGFCYSMLHKVSRFSALVIOQPV 60

QY 61 ELRNVACVFLVLRALDVTEDTSIPTDEKVPILIAFHRHIYDTHYSCGTKEYKILMD 120
DB 61 ELRDVACIFLVRALDVTEDTSIPTDVKVPILISPHQVHYDREWHFSCGTKEYKILMD 120

QY 121 QFHVSAFLELEKGYQEAIEEITRRMGAGMAKFCQEVETVDDYDEYCHYVAGLVGL 180
DB 121 QFHVSTAFLELRKHYYQQAIEDITMRMGAGMAKFCQEVETDDYDEYCHYVAGLVGL 180

QY 181 SKPLAAGSEVLTPDWEAISNMGFLQKTNIRDYLEDINEIPKSRMFWPRIWGYAD 240
DB 181 SKPLHASEKEDLASD--SLNSMGLFLQKTNIRDYLEDINEVPCRMFWPRIWTSYVN 238

QY 241 KLEDKYEENTKSVQCLNEMVTNALMHIEDCLKYMVSLRDPISFRCAIPQVMAIGTLA 300
DB 239 KLEELKEDTNSAKVQCLNEMVTNALPHVEDCLTYNGALRDPISFRCAIPQVMAIGTLA 298

QY 301 LCYNNEQVFRGVKRLRGTLAKVIDRTKTADVYGAFFDFSCMLKTKVDKNDPNASKTLN 360
DB 299 MCYDNIQEVFRGVKRLRGTLAKVIDRTKTADVYGAFFDFSCMLKSKVNNNDPNATKTLK 358

QY 361 RLEAVQKLCRDAGVLQNRKSV 382
DB 359 RLEVILKTCRDSGTLNKRKSVI 380

RESULT 3
ID_FDT_MOUSE STANDARD; PRT; 416 AA.
AC P53796;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Farnesyl-diphosphate farnesyltransferase (EC 2.5.1.21) (Squalene synthetase) (SQS) (SS) (FPP:PPP farnesyltransferase).
GN FDT1 OR ERG3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus; Mus.
OX NCBI_taxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6 X CBA; TISSUE=Liver;
RX MEDLINE=95092793; PubMed=7999794;
RA Inoue T., Osumi T., Hata S.;
RT "Molecular cloning and functional expression of a cDNA for mouse squalene synthase."
RL Biochim. Biophys. Acta 1260:49-54(1995).
CC -!- FUNCTION: CRITICAL BRANCH POINT ENZYME OF ISOPRENOID BIOSYNTHESIS THAT IS THOUGHT TO REGULATE THE FLUX OF ISOPRENE INTERMEDIATES THROUGH THE STEROL PATHWAY.
CC -!- CATALYTIC ACTIVITY: 2 farnesyl diphosphate = diphosphate + presqualene diphosphate.

```

```

CC -!- CATALYTIC ACTIVITY: Presqualene diphosphate + NADPH = diphosphate + squalene + NADP(+).
CC -!- COFACTOR: MAGNESIUM.
CC -!- PATHWAY: Critical branch point enzyme of isoprenoid and cholesterol biosynthesis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic reticulum.
CC -!- SIMILARITY: BELONGS TO THE PHYTOENE/SQUALENE SYNTHETASE FAMILY.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D29016; BAA06102.1; -.
DR PIR; S52075; S52075.
DR MGD; MGI:102706; Fdft1.
DR InterPro; IPR002060; Squ/phyt synthase.
DR Pfam; PF00494; SQS_PSY; 1.
DR TIGRFAMs; TIGR01559; squa1 synth; 1.
DR PROSITE; PS01044; SQUALEN_PHYTOEN_SYN_1; 1.
DR PROSITE; PS01045; SQUALEN_PHYTOEN_SYN_2; 1.
KW Multifunctional enzyme; Transferrase; Oxidoreductase; NADP; Magnesium;
KW Isoprene biosynthesis; Cholesterol biosynthesis; Transmembrane; Endoplasmic reticulum.
FT TRANSMEM 284 304
FT TRANSMEM 384 404
SQ SEQUENCE 416 AA; 48126 MW; FF653625227CF92A CRC64;

Query Match      41.3%; Score 848; DB 1; Length 416;
Best Local Similarity 46.9%; Pred. No. 4.3e-58;
Matches 168; Conservative 70; Mismatches 108; Indels 12; Gaps 5;

QY 8 LRYPDDDIYPLLMKRA-----IEKAEKQIPPEPHMGFCYSMLHKVSRFSLSVIQOINTEL 62
DB 7 LGHPEEFYLLRFRMGRRNFIPKMD-QDSLSLTKCYKLYNQTSRFAAVIQALDGD 65

QY 63 RNAVCVFLVLRALDVTEDTSIPTDEKVPILIAFHRHIYDTHYSCGTKEYKILMDQF 122
DB 66 RHAICVFLVLRALDVTEDDMSISVEKKIPLLCNFHTFLYDPEWRFTESKEKDRQVLEDF 125

QY 123 HHVSAFLELEKGYQEAIEEITRRMGAGMAKFCQEVETVDDYDEYCHYVAGLVGLSK 182
DB 126 PTISLEFRNLAEKYQTVVDDICHQMGCGMAEFVKDQVTSKQDMWKYCHYVAGLVGLSLR 185

QY 183 LFLAAGSE--VLTPDWEAISNMGFLQKTNIRDYLEDINEIPKSRMFWPRIWGYAD 240
DB 186 LFSASEFEDPIVGEDIETC-ANSNGLFLQKTNIRDYLEDQEE---GRKFWPQVWGYIK 241

QY 241 KLEDKYEENTKSVQCLNEMVTNALMHIEDCLKYMVSLRDPISFRCAIPQVMAIGTLA 300
DB 242 KLEDFAKPNVDVAVQCLNELITLQHIPDLVTLVSLRNLQSVNFENCAIPQVMAIATLA 301

QY 301 LCYNNEQVFRGVKRLRGTLAKVIDRTKTADVYGAFFDFSCMLKTKVDKNDPNASKT 358
DB 302 ACYNNOQVFGVKIRKQGVATLMDATNMPAVKAIYQVIEIYHRIIPNSDPSSSKT 359

RESULT 4
ID_FDT_RAT STANDARD; PRT; 416 AA.
AC Q02769;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Farnesyl-diphosphate farnesyltransferase (EC 2.5.1.21) (Squalene synthetase) (SQS) (SS) (FPP:PPP farnesyltransferase).
GN FDT1.
OS Rattus norvegicus (Rat).

```

| | |
|---------------------------|--|
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |
| OC | Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus. |
| OX | NCBI_TaxID=10116; |
| [1] | SEQUENCE FROM N.A. |
| RN | MEDLINE=9223095; PubMed=1569107; |
| RP | Shechter I., Klinger E., Rucker M.L., Engstrom R.G., Spirito J.A., |
| RA | Islam M.A., Boettcher B.R., Wienschein D.B.; |
| RT | "Solubilization, purification, and characterization of a truncated |
| RL | form of rat hepatic squalene synthetase."; |
| RT | J. Biol. Chem. 267:8628-8635(1992). |
| [2] | SEQUENCE FROM N.A. |
| RN | MEDLINE=93016074; PubMed=1400448; |
| RP | McKenzie T.B., Jiang G., Straubhaar J.R., Conrad D.G., Shechter I.; |
| RA | "Molecular cloning, expression, and characterization of the cDNA for |
| RT | the rat hepatic squalene synthase."; |
| RL | J. Biol. Chem. 267:21368-21374(1992). |
| [3] | SUBCELLULAR LOCATION. |
| RN | TISSUE=Liver; |
| RP | MEDLINE=9286129; PubMed=8509416; |
| RA | Stamellos K.D., Shackelford J.E., Schechter I., Jiang G., Conrad D.G., |
| RT | Keller G.-A., Krisans S.K.; |
| RL | "Subcellular localization of squalene synthase in rat hepatic cells. |
| RT | Biochemical and immunological evidence."; |
| RL | J. Biol. Chem. 268:12825-12836(1993). |
| CC | -!- FUNCTION: CRITICAL BRANCH POINT ENZYME OF ISOPRENOID BIOSYNTHESIS |
| CC | THAT IS THOUGHT TO REGULATE THE FLUX OF ISOPRENE INTERMEDIATES |
| CC | THROUGH THE STEROL PATHWAY. |
| CC | -!- CATALYTIC ACTIVITY: 2 farnesyl diphosphate = diphosphate + |
| CC | presqualene diphosphate. |
| CC | -!- CATALYTIC ACTIVITY: Presqualene diphosphate + NADPH = diphosphate |
| CC | + squalene + NADP(+). |
| CC | -!- COFACTOR: MAGNESIUM. |
| CC | -!- PATHWAY: Critical branch point enzyme of isoprenoid and |
| CC | cholesterol biosynthesis. |
| CC | -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic |
| CC | reticulum. |
| CC | -!- SIMILARITY: BELONGS TO THE PHYTOENE/SQUALENE SYNTHETASE FAMILY. |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation - |
| CC | the European Bioinformatics institute. There are no restrictions on its |
| CC | use by non-profit institutions as long as its content is in no way |
| CC | modified and this statement is not removed. Usage by and for commercial |
| CC | entities requires a license agreement (see http://www.isb-sib.ch/announce/ |
| CC | or send an email to license@isb-sib.ch). |
| CC | ----- |
| DR | EMBL; M95591; AAA42179.1; . |
| DR | FIR; A45105; A45105. |
| DR | InterPro; IPRO02060; Squ/phyt_synthes. |
| DR | InterPro; IPR006449; Squal_synth. |
| DR | Pfam; PF00494; SOS_PSY; 1. |
| DR | TIGRFAMs; TIGR01559; squal_synth; 1. |
| DR | PROSITE; PS01044; SQUALEN_PHYTOEN_SYN 1; 1. |
| DR | PROSITE; PS01045; SQUALEN_PHYTOEN_SYN 2; 1. |
| KW | Multifunctional enzyme; Transferase; Oxidoreductase; NADP; Magnesium; |
| KW | Isooprene biosynthesis; Cholesterol biosynthesis; Transmembrane; |
| KW | Endoplasmic reticulum. |
| FT | POTENTIAL. |
| FT | TRANSMEM 284 304 |
| FT | VARIANT 1 64 MISSING (IN TRUNCATED, ACTIVE FORM). |
| SQ | SEQUENCE 416 AA; 48106 MW; F4BC4D09CF72169 CRC64; |
| Query Match | 41.2%; Score 845; DB 1; Length 416; |
| Best Local Similarity | 46.5%; Pred. No. 7.3e-58; |
| Matches 166; Conservative | 71; Mismatches 108; Indels 12; Gaps 5; |
| YQ | 8 LRYPDIVPLLMKRA-----TEKAQKIPPEPHMGFCYSMLHKVRSFSLVIQLNTEL 62 |
| Db | 7 LGHPEFYNNLLFRMGRRNFIPKMDRN-SLSNSLKTCYKYLDQTSRFAAVIQALDGD 65 |

RC TISSUE=Lung, and Urinary bladder;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- CATALYTIC ACTIVITY: 2 farnesyl diphosphate = diphosphate +
presqualene diphosphate.
CC -!- CATALYTIC ACTIVITY: Presqualene diphosphate + NADPH = diphosphate
+ squalene + NADP(+).
CC -!- COFACTOR: MAGNESIUM.
CC -!- PATHWAY: Critical branch point enzyme of isoprenoid and
cholesterol biosynthesis
CC -!- SUBUNIT: Monomer (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
reticulum.
CC -!- SIMILARITY: BELONGS TO THE PHYTOENE/SQUALENE SYNTHETASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; L06070; AAA60582.1; -;
CC EMBL; L06105; AAA36645.1; -;
CC EMBL; X69141; CAA48896.1; -;
CC EMBL; S76822; AAB33404.1; -;
CC EMBL; BC003573; AAH03573.1; -;
CC EMBL; BC029641; AAH29641.1; -;
CC PIR; A45998; A45998.
CC PDB; 1ZEF; 18-OCT-00.
CC Genew; HGNC:3629; PDFT1.
CC MIM; 184420; -;
CC GO; GO:0016021; C: integral to membrane; TAS.
CC GO; GO:0006694; P: steroid biosynthesis; TAS.
CC InterPro; IPR002060; Squ/phyt synthase.
CC InterPro; IPR006449; Squa_synth.
CC Pfam; PF00494; SQS_PSY; 1;
CC TIGRFAMs; TIGR01559; squal synth; 1.
CC PROSITE; PS01044; SQUALENE_PHYTOEN_SYN_1; 1.
CC PROSITE; PS01045; SQUALENE_PHYTOEN_SYN_2; 1.
CC Multifunctional enzyme; Transferrase; Oxidoreductase; NADP; Magnesium;
KW Isoprene biosynthesis; Cholesterol biosynthesis; Transmembrane;
KW Endoplasmic reticulum; Polymorphism; 3D-structure.
FT TRANSMEM 284 304 POTENTIAL.
FT TRANSMEM 384 404 POTENTIAL.
FT VARIANT 45 45
FT
FT VARIANT 392 392
FT
FT VARIANT 353 353
FT CONFLICT 402 402 D -> N (IN REF. 4).
FT CONFLICT 402 402 T -> A (IN REF. 3).
SQ SEQUENCE 417 AA; 48115 MW; D36CB8C832F827EC CRC64;

Query Match 40.0%; Score 820; DB 1; Length 417;
Best Local Similarity 45.2%; Pred. No. 6.2e-56;
Matches 165; Conservative 69; Mismatches 121; Indels 10; Gaps 4;

Qy 8 LRYPDIDYPLKLM-----KRAIEKAEKQIPPEPHWGFYCYMLHKVSRFSFSLVIOQLNTELR 63
Db 7 LGHPPEEYLVNVRFRIGGKRRKPKMDQDSSLSLKTCTCYKYNLQTSRSFAAVIQALDGMER 66

Qy 64 NAVCVFLVLRALDTVEDDTSIPTDEKVPILIAFRRHIYDTHYSCGTKEYKILMDQPH 123
Db 67 NAVCIFVLVRALDTLEDDMTISVERKVPILHNFHSLFYQDPWRFMSKSKDRVLDFDP 126

Qy 124 HVSAAFLLEKGYQEAIEETRRMGAGMAKFIQOEVETVDDYDYCHYVAGLVGLGSLK 183
Db 127 TISLEPRNLAEKQTVIADICRRMGIGMAEFLDKGHTVSEQEWDKYCHYVAGLVGLGSLR 186

Qy 184 FLAAGSE--VLTPDWEAISNMGFLQKTNIIIDYLEDINEIPKSRFWFRPREHWKGVADK 241
Db 187 FSASEFEDPLVGDTER-ANSMGLFLQKTNIIIDYLEDQO---GGREFWFPQEVMSRYVK 242

Qy 242 LEDLKYEENTKSVQCLNEMVTNALMHIEDCLKYMVSLRDPISFRFCAIPOIMAIGTLAL 301
Db 243 LGDFAPKPDIDLAVQCLNELITNALHHIPDVITYLSRLRQSVNFCAIPQVMAIATLAA 302

Qy 302 CYNNEQVFRGVKLRGLTKAVIDRTKTMADVYGAFYDFSCMLKTKYDKNDPNASKTLNR 361
Db 303 CYNQVQVFKGAVKIRKQAVTLMDATNMPAVKAIYQYMEEIYHRIIPSDSPSSSKTRQI 362

Qy 362 LEAVQ 366
Db 363 ISTR 367

RESULT 6
ID -!- PDFT CANAL STANDARD; PRT; 448 AA.
AC P78589;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Farnesyl-diphosphate farnesyltransferase (EC 2.5.1.21) (Squalene
synthetase) (SQS) (SS) (PPP:PPP farnesyltransferase).
GN ERG9.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IFO 1060;
RA Ishii N., Arisawa M., Aoki Y.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: 2 farnesyl diphosphate = diphosphate +
presqualene diphosphate.
CC -!- CATALYTIC ACTIVITY: Presqualene diphosphate + NADPH = diphosphate
+ squalene + NADP(+).
CC -!- COFACTOR: MAGNESIUM.
CC -!- PATHWAY: Critical branch point enzyme of isoprenoid and
cholesterol biosynthesis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
reticulum (By similarity).
CC -!- SIMILARITY: BELONGS TO THE PHYTOENE/SQUALENE SYNTHETASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; D89610; BAA13995.1; -;

Thu Nov 13 10:01:45 2003

```

DR InterPro; IPR002060; Squ/phyt synthase.
DR InterPro; IPR006449; Squ_synth.
DR Pfam; PF00494; SQS_PSY; 1.
DR TIGRFAMs; TIGR01559; sqal synth; 1.
DR PROSITE; PS01044; SQUALEN_PHYTOEN_SYN_1; 1.
DR PROSITE; PS01045; SQUALEN_PHYTOEN_SYN_2; 1.
DR Multifunctional enzyme; Transferase; Oxidoreductase; NADPH; Magnesium;
KW Isoprene biosynthesis; Cholesterol biosynthesis; Transmembrane;
KW Endoplasmic reticulum.
FT TRANSMEM 295 315
FT TRANSMEM 428 448
FT TRANSMEM 448 448
SQ SEQUENCE 448 AA; 5171 MW; 357C1LE2918E4863 CRC64;

Query Match 36.4%; Score 747; DB 1; Length 448;
Best Local Similarity 41.2%; Pred. No. 2.9e-50;
Matches 159; Conservative 78; Mismatches 127; Indels 22; Gaps 8;

QY 1 MGSGLTMLRYDDIYPLLMK---KRAIEKAERQIPPEHMGFCYSMLHKVSRPSFLVIOQ 57
DB 1 MGKFLQLLSHPTLKAIVQLFGFQPLHPGRDVRNDKELGRCYELLNLTSSRFAVIEE 59

QY 58 LNTLERNVAVCYLVLRALDTVEDDTSIPTDEKVPILIAFHRHIYTDW---HYSCGTKE 114
DB 60 LPELRDAWIVYLVLRALDTIEDDMTKSSIKIPLLREFDTKLNTKWTFDGYGPNEXD 119

QY 115 KYLMQDFHVSAAFLKQVQAEIEITRMGMAGMAKFCQE-----VETVDDYDEY 168
DB 120 RTVLV-EFDKILNVVHRLKPOQYDIKISITFMGMGMADYILDBEFNVYGVATVEDVNL 178

QY 169 CHVAVGLVGLSKFLAA--GSEVLTPDWEAISNMGFLQKTNIRDYLEDINEIPKS 226
DB 179 CHVAVGLVGEGLNLFVLANFGDKILTENNFAKADSMGLFLOKTNIRDYHEDLQD---G 235

QY 227 RFWFPEIRMGKYADKLEDLYEENTNK---SVOCLEWVYTNALMHIEDCLYKVMSLRDP 283
DB 236 RSWFPEIRMSKYTENLQDFHKVKTAKEPAGVSCINELVNLALGHVTDCLDYLVLVQDPS 295

QY 284 IFFCAIPQIMAGTILALCYNEQVPRGVKLRGLTAKVIDRTKTMDVYGAFFDFSCM 343
DB 296 SFSFCAIPQVMAVATLAEVYNNPKVLHGVVKIRKGTTCRLILLESRTLFGVVKIFKEYIQV 355

QY 344 LKTKVDKNDPNASKTLNRLVAVQKLC 369
DB 356 INHSSVRDPENYKIGIKGEIEQYC 381

RESULT 7
ID_FDPY YEAST STANDARD; PRT; 444 AA.
AC P29704; 1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 31, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Farnesyl-diphosphate farnesyltransferase (EC 2.5.1.21) (Squalene
DE synthetase) (SQS) (SS) (PPP:PPP farnesyltransferase).
GN ERG9 OR YHR190W
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92224278; PubMed=1807826;
RA Peguier M., Richard L., Charles A.D., Karst F.;
RT "Isolation and primary structure of the ERG9 gene of Saccharomyces.
RL Curr. Genet. 20:365-372(1991).
RL [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91296756; PubMed=2068081;
RA Jennings S.M., Tsay Y.H., Fisch T.M., Robinson G.W.;
RT "Molecular cloning and characterization of the yeast gene for
RT squalene synthetase.";
```

```

Proc. Natl. Acad. Sci. U.S.A. 88:6038-6042 (1991).
[3]
SEQUENCE FROM N.A.
STRAIN=S288c / AB972;
RX MEDLINE=94378003; PubMed=8091229;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
RA Du Z., Favella A., Fulton L., Gattung S., Geisel C., Kirsten J.,
RA Kucaba T., Hillier L., Jier M., Johnston L., Langston Y.,
RA Latreille P., Louis E.J., Macri C., Mardis E., Menezes S., Mouser L.,
RA Nhan M., Rifkin L., Riles L., St Peter H., Trevaskis E., Vaughan K.,
RA Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,
RA Vaudin M.;
RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
VIII."
Science 265:2077-2082(1994).
CC -!- FUNCTION: MAY REGULATE THE FLUX OF ISOPRENE INTERMEDIATES THROUGH
THE STEROL PATHWAY. SQUALENE SYNTHETASE IS CRUCIAL FOR BALANCING
THE INCORPORATION OF FARNESYL DIPHOSPHATE (FDP) INTO STEROL AND
NONSTEROL ISOPRENE SYNTHESIS. ERG9 IS ALSO ESSENTIAL FOR CELL
GROWTH IN YEAST.
CC -!- CATALYTIC ACTIVITY: 2 farnesyl diphosphate = diphosphate +
presqualene diphosphate.
CC -!- CATALYTIC ACTIVITY: Presqualene diphosphate + NADPH = diphosphate
+ squalene + NADP(+).
CC -!- COFACTOR: MAGNESIUM.
CC -!- PATHWAY: Critical branch point enzyme of isoprenoid and
cholesterol biosynthesis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
reticulum (by similarity).
CC -!- SIMILARITY: BELONGS TO THE PHYTOENE/SQUALENE SYNTHETASE FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
-----
EMBL; M63979; AAA34597.1; -
EMBL; X59959; CAA42583.1; -
EMBL; U00030; BAB68360.1; -
PIR; S46682; S46682.
DR SGD; S0001233; ERG9.
DR InterPro; IPR002060; Squ/phyt synthase.
DR InterPro; IPR006449; Squ_synth.
DR Pfam; PF00494; SQS_PSY; 1.
DR TIGRFAMs; TIGR01559; sqal synth; 1.
DR PROSITE; PS01044; SQUALEN_PHYTOEN_SYN_1; 1.
DR PROSITE; PS01045; SQUALEN_PHYTOEN_SYN_2; 1.
KW Multifunctional enzyme; Transferase; Oxidoreductase; NADPH; Magnesium;
KW Isoprene biosynthesis; Cholesterol biosynthesis; Transmembrane;
KW Endoplasmic reticulum.
FT TRANSMEM 291 311
FT TRANSMEM 421 441
FT TRANSMEM 441 441
FT CONFLICT 48 48 L -> F (IN REF. 1).
FT CONFLICT 286 286 G -> S (IN REF. 2).
FT CONFLICT 320 320 N -> D (IN REF. 1).
FT CONFLICT 330 330 Y -> C (IN REF. 1).
FT CONFLICT 429 429 L -> S (IN REF. 1).
SQ SEQUENCE 444 AA; 51719 MW; 3DA00295B8DFDB16 CRC64;

Query Match 35.2%; Score 721; DB 1; Length 444;
Best Local Similarity 41.0%; Pred. No. 2.9e-46;
Matches 157; Conservative 72; Mismatches 138; Indels 16; Gaps 6;

QY 1 MGSGLTMLRYDDIYPLLMK---KRAIEKAERQIPPEHMGFCYSMLHKVSRPSFLVIOQL 58
DB 1 MGKLLQLALHPVEMKAAKLFKFCRTPLFSIYDQSTSPVLLHCFELLNLTSSRFAVIREL 60

QY 59 NTELRNVCYFVLVLRALDTVEDDTSIPTDEKVPILIAFHRHIYTDWYHSCGTKEYK-- 116
DB 61 HPELRNCVTLFYLLRALDTIEDDMISIEHDLKILLRHFHEKLLLTKWSFGNAPDVKDR 120
```

QY 117 ILMDQFHVGAFLLEKQYQAEIETRRMGAGMAKFIQOE-----VETVDDYDYCH 170
 Db 121 AVLTDFESILIEFHKLKPEQVEIKETRMGNGMADYILDENYNLGLQTVHDYDYCH 180
 QY 171 YVAGLVGLGSLFLAA--GSEVLTPDWEAISNMGFLFLOKTNIRYDLEINEIPKSRM 228
 Db 181 YVAGLVGGTLLIVIAKFAINESYSN-EQYESMGFLFLOKTNIRYDNEVDVD---GRS 236
 QY 229 FWPKEIMGKADKLEDKYBENTNKSVOCLNEMVTNLMHIEDCLKYMSLRDPSIRPFC 288
 Db 237 FWPKEIMSQAPQLKFWKPENEQGLGDCINHLVLNLSHVIDVLTLAGIHEQSTQFC 296
 QY 289 AIPQIMAGTIALCYNNEQVGRGVVULRGTLAKVIDRTKTMADYGAFYDFSCMLTKV 348
 Db 297 AIPQVMAIATLAVFNRRVHLGNVKIRKGTTCVLIILKSRTRGCVEIFDYLRDIKSKL 356
 QY 349 DKNDPNASKTLNRLVAVOKLCRD 371
 Db 357 AVQDPNFKLNIQISKIEQFMEE 379

RESULT 8
 FDFT_CANGA STANDARD; PRT; 443 AA.
 AC Q9HG26;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DE Farnesyl-diphosphate farnesyltransferase (EC 2.5.1.21) (Squalene synthetase) (SQS) (SS) (PPP:PPP farnesyltransferase).
 GN ERG9.
 OS Candida glabrata (Yeast) (Torulopsis glabrata).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 OX NCBI_TaxID=5478;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=ATCC 2001;
 RX MEDLINE=20408567; PubMed=10952589;
 RA Nakayama H., Izuta M., Nakayama N., Arisawa M., Aoki Y.;
 RT "Depletion of the squalene synthase (ERG9) gene does not impair growth of Candida glabrata in mice."
 RL Antimicrob. Agents Chemother. 44:2411-2418(2000).
 CC -!- CATALYTIC ACTIVITY: 2 farnesyl diphosphate = diphosphate + presqualene diphosphate.
 CC -!- CATALYTIC ACTIVITY: Presqualene diphosphate + NADPH = diphosphate + squalene + NADP(+).
 CC -!- COFACTOR: Magnesium (By similarity).
 CC -!- PATHWAY: Critical branch point enzyme of isoprenoid and cholesterol biosynthesis.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic reticulum (By similarity).
 CC -!- SIMILARITY: BELONGS TO THE PHYTOENE/SQUALENE SYNTHETASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AB009978; BAB12207.1; -.
 DR InterPro: IPR002060; Squ/phyt_synthese.
 DR InterPro: IPR006449; Squa_synth.
 DR Pfam: PF00494; SQS_PSY; 1.
 DR TIGRFAMs: TIGR01559; squa1_synth; 1.
 DR PROSITE: PS01044; SQUALEN_PHYTOEN_SYN_1; 1.
 DR PROSITE: PS01045; SQUALEN_PHYTOEN_SYN_2; 1.
 KW Multifunctional enzyme; Transferase; Oxidoreductase; NADP; Magnesium;
 KW Isoprene biosynthesis; Cholesterol biosynthesis; Transmembrane;
 KW Endoplasmic reticulum.

FT TRANSMEM 181 201 POTENTIAL.
 FT TRANSMEM 294 314 POTENTIAL.
 SQ SEQUENCE 443 AA; 51379 MW; 4176FD68E2A0F284 CRC64;
 Query Match 34.6%; Score 709.5; DB 1; Length 443;
 Best Local Similarity 41.2%; Pred. No. 2.2e-47;
 Matches 157; Conservative 67; Mismatches 120; Indels 37; Gaps 7;
 QY 1 MGSGLTMRYPDDIYPLKMK-----RAIEKAERKQIPPEPHMGFCYSMLHKVS 48
 Db 1 MGKVLDDLALHPLLELRAALKLKFIRQPLFSTNDTRATPOLER-----CYELLNLS 50
 QY 49 RPSFLVIQQLNTELRNACVFIYLVLRALDVTVEDDTSIPTDEKVPILIAFHHRYDTHY 108
 Db 51 RSPAIVIMELHPELRNVIWFIYLVLRALDVTVEDDTSIPTDEKVPILIAFHHRYDTHY 110
 QY 109 SCGTKEVK--ILMDQFHVGAFLLEKQYQAEIETRRMGAGMAKFIQOE-----VE 160
 Db 111 DGNLKEKORVLTPEPCILGEYHKUKPEYQKVIKRTGLMGNGMADYILDENFNMGVQ 170
 QY 161 TVDDYDEYCHYVAGLVGLGSLFLAA--GSEVLTPDWEA--ISNSMGLFLQKTNIRYD 216
 Db 171 TVKDYDKYCHYVAGLVGDGLTELIVLAGFGSDDLHYGKNSFQLYESLGLFLOKTNIRYD 230
 QY 217 LEDINEIPKSRMPWPRIENGKYADKLEDLYKENTNKSVOCLNEMVTNLMHIEDCLKYM 276
 Db 231 AEDLDD---GRSEWPKKEINSEYATKLTDFRDPKNTQKGVDCINHLVLNALTHTVIDVLT 287
 QY 277 VSLRDPISIRFCAIPQIMAGTIALCYNNEQVGRGVVULRGTLAKVIDRTKTMADYGA 336
 Db 288 SSIHEOSSPQCAIPQVMAIATLAVFNRRVHLGNVKIRKGTTCVLIILKSRTRGCV 347
 QY 337 FYDFCMLTKVDKNDPNASK 357
 Db 348 FQYVLRDMKQRLPVEDPNYLK 368

RESULT 9
 FDFT_SCHPO STANDARD; PRT; 460 AA.
 AC P36536;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Farnesyl-diphosphate farnesyltransferase (EC 2.5.1.21) (Squalene synthetase) (SQS) (SS) (PPP:PPP farnesyltransferase).
 GN ERG9 OR SPB646.05C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=93233634; PubMed=8474436;
 RA Robinson G.W., Tsay Y.H., Kienzie B.K., Smith-Monroy C.A., Bishop R.W.;
 RT "Conservation between human and fungal squalene synthetases: similarities in structure, function, and regulation."
 RL Mol. Cell. Biol. 13:2706-2727(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A., Scouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S., Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M., Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A., Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G., Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K., James K., Jones L., Jones M., Leather S., McDonald S., McLean J., Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,

| | | | | | |
|----|--|----|-----|---|-----|
| RA | Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E., Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S., Skelton J., Simmonds M., Squares R., Squares S., Stevens K., Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S., Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B., Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H., Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M., Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B., Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S., Gaubert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M., Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G., Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J., Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L., Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J., Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.; "The genome sequence of Schizosaccharomyces pombe."; | DB | 216 | QKVNIIIRDYREDDDD---NRHFWPREIWSKYTSFEGDLCLPDNSEKALECISDMTANALT | 272 |
| CC | Nature 415:871-880(2002). | CC | 328 | KTMADVYGAFYDFSCMLKTKVDKNDPNASK | 357 |
| CC | -!- CATALYTIC ACTIVITY: 2 farnesyl diphosphate = diphosphate + presqualene diphosphate. | CC | 333 | VNLKNVCDLFLRYTRDIHYKNTPKDPNFK | 362 |
| CC | -!- CATALYTIC ACTIVITY: Presqualene diphosphate + NADPH = diphosphate + squalene + NADP(+). | CC | | | |
| CC | -!- COFACTOR: MAGNESIUM. | CC | | | |
| CC | -!- PATHWAY: Critical branch point enzyme of isoprenoid and cholesterol biosynthesis. | CC | | | |
| CC | -!- SUBUNIT: Monomer (By similarity). | CC | | | |
| CC | -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic reticulum (By similarity). | CC | | | |
| CC | -!- SIMILARITY: BELONGS TO THE PHYTOENE/SQUALENE SYNTHETASE FAMILY. | CC | | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch). | CC | | | |
| CC | EMBL; L06071; AAA35343.1; -- | CC | | | |
| DR | EMBL; AL035216; CAA22809.1; -- | CC | | | |
| DR | PIR; T40581; T40581. | CC | | | |
| DR | GeneDB SPombe; SPBC646.05c; -- | CC | | | |
| DR | InterPro; IPR002060; Squ/phyt synthase. | CC | | | |
| DR | InterPro; IPR006449; Squal_synth. | CC | | | |
| DR | Pfam; PF00494; SQS_PSY; 1. | CC | | | |
| DR | TIGRFAMS; TIGR01559; squal_synth; 1. | CC | | | |
| DR | PROSITE; PS01044; SQUALEN_PHYTOEN_SYN_1; 1. | CC | | | |
| DR | PROSITE; PS01045; SQUALEN_PHYTOEN_SYN_2; 1. | CC | | | |
| KW | Multifunctional enzyme; Transferase; Oxidoreductase; NADP; Magnesium; | CC | | | |
| KW | Isoprene biosynthesis; Cholesterol biosynthesis; Transmembrane; | CC | | | |
| KW | Endoplasmic reticulum. | CC | | | |
| FT | TRANSMEM 288 308 POTENTIAL. | CC | | | |
| FT | TRANSMEM 425 445 POTENTIAL. | CC | | | |
| FT | SEQUENCE 460 AA; 53321 MW; BAEF7F24B6CE5D25 CRC64; | CC | | | |
| CC | | CC | | | |
| CC | Query Match 34.2%; Score 700.5; DB 1; Length 460; | CC | | | |
| CC | Best Local Similarity 44.5%; Pred. No. 1.1e-46; | CC | | | |
| CC | Matches 147; Conservative 60; Mismatches 106; Indels 17; Gaps 5; | CC | | | |
| CC | | CC | | | |
| QY | 40 CYSMHLKVSFSLVIOQLNTELRNAVCFYLVLRALDVTDEDDTSIPTDEKVPILIAFHR 99 | QY | 40 | CYSMLHKVSRSFSLVIOQLNTELRNAVCFYLVLRALDVTDEDDTSIPTDEKVPILIAFHR 99 | |
| DB | 38 CYQLLDMTSSFAVVIKELPENGIREAVMIFFYLVLRGLDVTDEDDMTLPDKKLPILRDFYK 97 | DB | 38 | CYQLLDMTSSFAVVIKELKGDVSRVICLFYLVLRALDVTDEDDMTIPQRIPLLVNFYK 109 | |
| QY | 100 HIYTDWHY--SCGTKEYKILMPQFHVSAFAFLEKGYQEAIEETTRMGAGMAKFIQ 157 | QY | 100 | HIYTDWHY--SCGTKEYKILMDQPHVSAFAFLEKGYQEAIEETTRMGAGMAKFI-- 155 | |
| DB | 98 TIEVEGTFNMSGNEKDRQLLVFDDVVIKEYLNLSEGYRNVISNITKMGDGMAYVASL 157 | DB | 98 | TIEVEGTFNMSGNEKDRQLLVFDDVVIKEYLNLSEGYRNVISNITKMGDGMAYVASL 157 | |
| QY | 158 -----EVEVDYDYCHYVAGLVLGSLKFLAAGSEVLTPDW---EATSNMGLFL 207 | QY | 158 | -----EVEVDYDYCHYVAGLVLGSLKFLAAGSEVLTPDW---EATSNMGLFL 207 | |
| DB | 158 AEKNDGFSVETIEDFNKYCHYVAGLVIGLSRLF--AQSKLDPDLAHSQATISLGLFL 215 | DB | 158 | AEKNDGFSVETIEDFNKYCHYVAGLVIGLSRLF--AQSKLDPDLAHSQATISLGLFL 215 | |
| QY | 208 OKTNIIIRDYLEDINEIPKSMFMEITWGVADKLEDLKYEENTNKSQVCLNEMVNTALM 267 | QY | 208 | OKTNIIIRDYLEDINEIPKSMFMEITWGVADKLEDLKYEENTNKSQVCLNEMVNTALM 267 | |

-1- SUBUNIT: Monomer (By similarity).
-1- SUBCELLULAR LOCATION: Chloroplast.
-1- SIMILARITY: BELONGS TO THE PHYTOENE/SQUALENE SYNTHETASE FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch)

EMBL: U32636; AAB60314.1; -;
PIR: S68307; S68307.
MaizeDB: 66643; 66643.
InterPro: IPR002060; Squ/phyt_synthese.
Pfam: PF00494; SQS_PSY; 1.
PROSITE: PS01044; SQUALEN_PHYTOEN_SYN_1; 1.
PROSITE: PS01045; SQUALEN_PHYTOEN_SYN_2; 1.
MultiFunctional enzyme; Carotenoid biosynthesis;
Isoprene biosynthesis; Transferase; Chloroplast; Transit peptide.
TRANSIT ? CHLOROPLAST (POTENTIAL).
CHAIN ? 410 PHYTOENE SYNTHASE.
VARIANT 344 344 N -- T (IN ALLELE B73).
SEQUENCE 410 AA; 46481 MW; 21070A33624EED79 CRC64;

| | | | | |
|-----------------------|--------|------------------|-----------------|---------------------|
| Query Match | 5.8%; | Score 118; | DB 1; | Length 410; |
| Best Local Similarity | 22.3%; | Pred. No. 0.08; | | |
| Matches | 64; | Conservative 47; | Mismatches 118; | Indels 58; Gaps 13; |

| | | | |
|-----|-----|--|-----|
| QY | 14 | IYLLKMKRAIEKAKEIQIP-----PE-----PHWGF-----CYSMLHKVKSFSFSLVIQQL | 58 |
| Ddb | 81 | VVDVLQAALLKRLRTFPVLDARPQDMDPNGLKEAYDRGCEIECYAKTFYLGTMLM | 140 |
| QY | 59 | NTELRNAVCFYLVLRALDVTVEDDTs----IPTDEKVPIIAFHRIHYIDTWHSYCGTKE | 114 |
| Ddb | 141 | TEERRAIWAIYWCRRTDELVDGNANYITPT-----ALDR-----WE----- | 179 |
| QY | 115 | YKLMQDPHHVSAAFLBELEGYOEAIEEIT-----RRMGAGNAKFI COEVTVDDYDE- | 167 |
| Ddb | 180 | KRLIEDLF--TGPRPYMDLAALSDTISRFPIDIQPFPRDMI EGMRSDL--RKTRYNNFDEL | 234 |
| QY | 168 | --YCHYVAGLVGL--GLSKFLAAGESEVLTPDWEAINSMGLFLQKTIIRI DYLEDINEIP | 224 |
| Ddb | 235 | WMYCIYVAGTVGLMSVPWGIAIESKATTESVYSAAALGANQLTNILRDVGEDAR--- | 291 |
| QY | 225 | KSRMPWPREGWKYADLKLEDUKYEENTKNKSVOCLNEMVNVMALMHIED | 271 |
| Ddb | 292 | RGRIVLPQDELAQGLSDIEDIFKGVTNRWRNFMRKQIKRARMFFEE | 338 |

RESULT 13
PSY DAUCA
ID PSY DAUCA STANDARD; PRT; 398 AA.
AC Q9SSU8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phytoene synthase, chloroplast precursor [EC 2.5.1.-].
DN PSY.
OS Daucus carota (Carrot).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; campanulids; Apiales; Apiaceae; Daucus.
OX NCBI_TaxID=4039;
RN [1]
RS SEQUENCE FROM N.A.
RC RP
RC TISSUE=Root;
RA Uno T.; Sankawa U.;
RT "Daucus carota phytoene synthase."
RL Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
CC -!- FUNCTION: CATALYZES THE REACTION FROM PREPHYTOENE DIPHOSPHATE TO PHYTOENE

Nakamura K., Harashima K.;
"Elucidation of the Erwinia uredovora carotenoid biosynthetic pathway


```
RT by functional analysis of gene products expressed in Escherichia
RL coli."
RN J. Bacteriol. 172:6704-6712(1990).
RP [2]
RA REVISIONS TO N-TERMINUS.
RA Misawa N, Nakagawa M., Kobayashi K., Yamano S., Izawa Y.,
RA Nakamura K., Harashima K.;
RA Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CATALYZES THE REACTION FROM PREPHYTOENE DIPHOSPHATE
CC TO PHYTOENE.
CC -1- CATALYTIC ACTIVITY: 2 geranylgeranyl diphosphate = diphosphate +
CC prephytoene diphosphate.
CC -1- CATALYTIC ACTIVITY: Prephytoene diphosphate = diphosphate +
CC phytoene.
CC -1- SIMILARITY: BELONGS TO THE PHYTOENE/SQUALENE SYNTHETASE FAMILY.
CC -1- SIMILARITY: SOME, TO Y4AD.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D90087; BAA14128.2; -.
CC PIR: E37802; E37802.
CC InterPro: IPR002060; Squ/phyt_synthes.
CC Pfam: PF00494; SQS_PSY; 1.
CC PROSITE: PS01044; SQUALEN_PHYTOEN_SYN_1; 1.
CC PROSITE: PS01045; SQUALEN_PHYTOEN_SYN_2; 1.
CC Multifunctional enzyme; Carotenoid biosynthesis; Transferase.
CC SEQUENCE 309 AA; 34472 MW; 9AA381A7376BBFC9 CRC64;
CC -----
Query Match 5.6%; Score 115; DB 1; Length 309;
Best Local Similarity 21.8%; Pred. No. 0.094;
Matches 51; Conservative 43; Mismatches 102; Indels 38; Gaps 9;
QY 48 SRSLSVIQQLNTELNAVCFYLVLRALDVTEDDTSIPTDEKVPILIAFHRHIYDTDW - 106
DB 18 SKSFATSKLFDKATRRSVLMVWCHDCHDDVDDQTLGFGAQQPALQTPQRLMQLEMK 77
QY 107 --HYSCGTEKYKILMDOFHVSA-----AFLELEKGYQAEIEITRRMGAGMAKFC 156
DB 78 TRQAYAGSOMHEPAFAFQVAMAHDIAPAFDHLF-GFAMDVRE----- 122
QY 157 QEVETVDDYDEYCHYVAGLVGLGSLKFLAAGSEVLTDPWEAISNSMGLFLOKTNIRDY 216
DB 123 AQYSQDDTLRYCVHVGAVVGLMMAQIMGVDRDNLDR-----ACDLGLAFQLTNIARDI 177
QY 217 LEDINEIPKSRMFWPREIMGKYADKLEDLKYENTNKSVOCLNEMVTNALMHIE 270
DB 178 VDDAH---AGRCYLPAS-WLEH-EGLNKENYAAPENR--QALSRIARRLVQEA 224
RESULT 15
Y4AC RHISN STANDARD; PRT; 279 AA.
AC P55350;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Putative phytoene synthase (EC 2.5.1.-).
GN Y4AC.
OS Rhizobium sp. (strain NGR234).
OG Plasmid sym pNGR234a.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
OX NCBI_TaxID=394;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97305956; PubMed=9163424;
RA Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,
```

```
RA Perret X.;
RT "Molecular basis of symbiosis between Rhizobium and legumes.";
RL Nature 387:394-401(1997).
CC -1- FUNCTION: CATALYZES THE REACTION FROM PREPHYTOENE DIPHOSPHATE TO
CC PHYTOENE. PROBABLY PART OF AN OPERON Y4AABCD INVOLVED IN THE
CC SYNTHESIS OF AN ISOPRENOID COMPOUND.
CC -1- CATALYTIC ACTIVITY: 2 geranylgeranyl diphosphate = diphosphate +
CC prephytoene diphosphate.
CC -1- CATALYTIC ACTIVITY: Prephytoene diphosphate = diphosphate +
CC phytoene.
CC -1- SIMILARITY: BELONGS TO THE PHYTOENE/SQUALENE SYNTHETASE FAMILY.
CC -1- SIMILARITY: SOME, TO Y4AD.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: A500064; AAB91601.1; -.
CC InterPro: IPR002060; Squ/phyt_synthes.
CC Pfam: PF00494; SQS_PSY; 1.
CC PROSITE: PS01044; SQUALEN_PHYTOEN_SYN_1; 1.
CC PROSITE: PS01045; SQUALEN_PHYTOEN_SYN_2; 1.
CC Hypothetical protein; Multifunctional enzyme; Carotenoid biosynthesis;
CC Isoprene biosynthesis; Transferase; Plasmid.
CC SEQUENCE 279 AA; 31354 MW; DD236CB8F72BAFC5 CRC64;
CC -----
Query Match 5.5%; Score 112.5; DB 1; Length 279;
Best Local Similarity 26.2%; Pred. No. 0.13;
Matches 51; Conservative 26; Mismatches 75; Indels 43; Gaps 8;
QY 50 SRSLSVIQQLNTELNAVCFYLVLRALDVTEDDTSIPTDEKVPILIAFHRHIYDTDWHS 109
DB 19 SPYLGMRTELPVQREAIQIYSCFRCQVDDIADSD-PRHRLAALQQRDHI---DALYQ 74
QY 110 C----GTEKYKILMDOFHVSAFLELEKGYQ-EAIEEITRRMGAGMAKFCQEVETVDD 164
DB 75 CVPPPLKDYLASVTTFGLKREDFLAVDGMMDVLQDIRAPKMYL----- 121
QY 165 YDEYCHYVAGLVG-----LGSLKFLAAGSEVLTDPWEAISNSMGLFLOKTNIRDY 218
DB 122 -DLYCDRVASVGRMSVRVFGLS-----EDGIALAHHLGRALQTLTILRDIDE 169
QY 219 DINEIPKSRMFWPRE 233
DB 170 DAG---LGRLYTPRE 181
```

Search completed: November 12, 2003, 16:29:31
Job time : 18 secs

| Result No. | Query | | | DB | ID | Description |
|------------|--------|-------|-----|----|--------|----------------------|
| | Score | Match | % | | | |
| 1 | 1702.5 | 83.0 | 413 | 10 | Q22107 | 022107 glycine max |
| 2 | 1697 | 82.7 | 413 | 10 | Q23118 | 023118 arabidopsis |
| 3 | 1692 | 82.5 | 413 | 10 | Q65868 | 065868 arabidopsis |
| 4 | 1671 | 81.5 | 413 | 10 | Q42760 | 042760 glycyrrhiza |
| 5 | 1663 | 81.1 | 415 | 10 | Q48666 | 048666 panax ginseng |
| 6 | 1659 | 80.9 | 412 | 10 | Q42761 | 042761 glycyrrhiza |
| 7 | 1638 | 79.9 | 411 | 10 | Q9XJ31 | 09XJ31 solanum tub |
| 8 | 1637 | 79.8 | 411 | 10 | Q24148 | 024148 nicotiana t |
| 9 | 1632 | 79.6 | 413 | 10 | Q8GS16 | 08GS16 medicago tr |
| 10 | 1629.5 | 79.4 | 411 | 10 | Q9XF02 | 09XF02 capsicum an |
| 11 | 1605 | 78.3 | 418 | 10 | Q9FUK1 | 09FUK1 artemisia a |
| 12 | 1446 | 70.5 | 401 | 10 | Q22106 | 022106 zea mays (m |
| 13 | 1437.5 | 70.1 | 376 | 10 | Q40472 | 040472 nicotiana t |
| 14 | 1426 | 69.5 | 403 | 10 | Q22105 | 022105 oryza sativ |
| 15 | 1258 | 61.3 | 290 | 10 | Q8LPN2 | 08LPN2 nicotiana t |
| 16 | 917.5 | 44.7 | 461 | 10 | Q9SDW9 | 09SDW9 botryococcu |

Db 181 SKFIASELEILTDPWKQISNSTGLFLOKTNIIKDYLEINERPKSMFWRPREIWKYVD 240
Qy 241 KLEDLKYEENTNKSVOCLNEMVTNLMHIEDCLKYMVSLRDPISIFRCAIPQIMAGTILA 300
Db 241 KLEDFKNEEKATKAVQCLNEMVTNLMHVEDCLKSLASLRDPAIFQSCAIPQIVAGTILT 300
Qy 301 LCYNNEQVFRGVVWKLRLGLTAKVIDRTKTMDVYGAFFDFSCMLTKTKVNDKNDPNAKTILN 360
Db 301 LCYNNEQVFRGVVWKLRLGLTAKVIDRTKTMDVYGAFFDFSCMLTKTKVNDKNDPNAKTILN 360
Qy 361 RLEAVQKLCRDAGVLQNRKSYVNDKQ 387
Db 361 RLEAVQKLCRDAGVLQNRKSYVNDKQ 387
PRT: 413 AA.
PRELIMINARY;
ID O65688
AC O65688;
DT 01-AUG-1998 (TremBLrel. 07, Created)
DT 01-NOV-1999 (TremBLrel. 12, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE Putative squalene synthase.
GN T4L20.230 OR Ar4G34650.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroside II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Terry N., Ardiles W., Buysshaert C., Dasseville R.,
RA De Clerck R., De Keyser A., Neyt P., Rouze P., Van Den Daele H.,
RA Villarroel R., Gielen J., van Montagu M., Jesse T., Heijnen L., Vos P.,
RA Hohnel J., Mewes H.W., Mayer K.F.X., Lemcke K., Schueller C.,
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Terry N., Ardiles W., Buysshaert C., Dasseville R., De Clerck R.,
RA De Keyser A., Neyt P., Rouze P., Van Den Daele H., Villarroel R.,
RA Gielen J., Van Montagu M., Mewes H.W., Lemcke K., Mayer K.F.X.,
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL023094; CAA18844.2; -
DR EMBL; AL161585; CAB80182.1; -
DR InterPro; IPR001064; Crystallin.
DR InterPro; IPR002060; Squ/phyt_synthse.
DR InterPro; IPR006449; Squal_synth.
DR Pfam; PF00494; SQS_PSY; 1.
DR TIGRFAMs; TIGR01559; squal_synth; 1.
DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; 1.
SQ SEQUENCE 413 AA; 47183 MW; 6F461C41D3849C15 CRC64;
Query Match 82.5%; Score 1692; DB 10; Length 413;
Best Local Similarity 80.6%; Pred. No. 2.3e-132;
Matches 312; Conservative 37; Mismatches 38; Indels 0; Gaps 0;
Qy 1 MGSGLTWMRYPDIIYPLLKMKRAIEKAEKQIPPEPHWGCYMLHKVSRFSLSVIOQLNT 60
Db 1 MGSGLTWMRYPDIIYPLLKMKRAIEKAEKQIPPEPHWGCYMLHKVSRFSLSVIOQLNT 60
Qy 61 ELRNVACVFLVLRALDTVEDDTSIPTDEKVPILIAFHRIYDTHWYSCGKEYKILMD 120
Db 61 ELRNVACVFLVLRALDTVEDDTSIPTDEKVPILIAFHRIYDTHWYSCGKEYKILMD 120
Qy 121 QFHVSAFALEKGYQEAIEETIRRMAGMAKFI COEVEVDYDYDEYCHYVAGLVGL 180
Db 121 QFHVSAFALEKGYQEAIEETIRRMAGMAKFI COEVEVDYDYDEYCHYVAGLVGL 180
Qy 181 SKFLAAGSEVLTPDWEAISNSMGLFLOKTNIIIRDYLEINERPKSMFWRPREIWKYAD 240

Db 61 ELRNVACVFLVLRALDTVEDDTSIPTDEKVPILIAFHRIYDTHWYSCGKEYKILMD 120
Qy 121 QFHVSAFALEKGYQEAIEETIRRMAGMAKFI COEVEVDYDYDEYCHYVAGLVGL 180
Db 121 QFHVSAFALEKGYQEAIEETIRRMAGMAKFI COEVEVDYDYDEYCHYVAGLVGL 180
Qy 181 SKFLAAGSEVLTPDWEAISNSMGLFLOKTNIIIRDYLEINERPKSMFWRPREIWKYAD 240
Db 181 SKFLAAGSEVLTPDWEAISNSMGLFLOKTNIIIRDYLEINERPKSMFWRPREIWKYAD 240
PRT: 413 AA.
PRELIMINARY;
ID O23118
AC O23118;
DT 01-JAN-1998 (TremBLrel. 05, Created)
DT 01-JAN-1998 (TremBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE Squalene synthase 2.
GN SQS2.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroside II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Kribbi R., Arro M., Del Arco A., Gonzalez V., Balcells L.I.,
RA Delorme D., Ferrer A., Karst F., Boron A.,
RT "Cloning and Characterization of the Arabidopsis thaliana SQS1 Gene
RT Encoding Squalene Synthase. Involvement of the C-Terminal Region of
RT the Enzyme in the Channeling of Squalene Through the Sterol Pathway.";
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF004396; AAB61927.1; -
DR InterPro; IPR001064; Crystallin.
DR InterPro; IPR002060; Squ/phyt_synthse.
DR InterPro; IPR006449; Squal_synth.
DR Pfam; PF00494; SQS_PSY; 1.
DR TIGRFAMs; TIGR01559; squal_synth; 1.
DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; 1.
SQ SEQUENCE 413 AA; 47306 MW; 0D524245AD067FF4 CRC64;
Query Match 82.7%; Score 1697; DB 10; Length 413;
Best Local Similarity 80.6%; Pred. No. 9e-133;
Matches 312; Conservative 38; Mismatches 37; Indels 0; Gaps 0;
Qy 1 MGSGLTWMRYPDIIYPLLKMKRAIEKAEKQIPPEPHWGCYMLHKVSRFSLSVIOQLNT 60
Db 1 MGSGLTWMRYPDIIYPLLKMKRAIEKAEKQIPPEPHWGCYMLHKVSRFSLSVIOQLNT 60
Qy 61 ELRNVACVFLVLRALDTVEDDTSIPTDEKVPILIAFHRIYDTHWYSCGKEYKILMD 120
Db 61 ELRNVACVFLVLRALDTVEDDTSIPTDEKVPILIAFHRIYDTHWYSCGKEYKILMD 120
Qy 121 QFHVSAFALEKGYQEAIEETIRRMAGMAKFI COEVEVDYDYDEYCHYVAGLVGL 180
Db 121 QFHVSAFALEKGYQEAIEETIRRMAGMAKFI COEVEVDYDYDEYCHYVAGLVGL 180
Qy 181 SKFLAAGSEVLTPDWEAISNSMGLFLOKTNIIIRDYLEINERPKSMFWRPREIWKYAD 240

Db 181 SKIFIASEILTPDWKQNSNSTGLFQKTNIIKDYLEDINERPKSRMFPRWIKGYVD 240
Qy 241 KLEDLKYEENTNKSVOCLNEMVTNALMHIEDCLKYMVSRLDPSIFRFAIPQIMAGTGLA 300
Db 241 KLEDFKNEEKATKAVOCLNEMVTNALMHVEDCLKSLASLRDPAIFOSCAIPQIIVAGTGLA 300
Qy 301 LCYNNEQVFRGVVGLRGLTAKVIDRTKTMADYVGFYDFSCMLTKYKVDKNDPNASKTLN 360
Db 301 LCYNNEQVFRGVVGLRGLTAKVIDRTKTMADYVGFYDFSCMLTKYKVDKNDPNASKTLN 360
Qy 361 RLEAVOKLCRDAGVLQNRKSYNDKGO 387
Db 361 RLETIKKCKENGGLHKRKSYNDETQ 387

RESULT 4

Q42760 PRELIMINARY; PRT; 413 AA.
AC Q42760;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Squalene synthase (EC 2.5.1.21).
OS Glycyrrhiza glabra.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Galegeae; Glycyrrhiza.
OX NCBI_TaxID=49827;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97070595; PubMed=8913521;
RA Hayashi H., Hiraoka N., Ikeshiro Y.;
RT "Molecular cloning and functional expression of cDNAs for glycyrrhiza
RL glabra squalene synthase.";
RL ENBL; D86409; BAA13083.1; -;
DR InterPro; IPR002060; Squ/phyt_synthase.
DR InterPro; IPR006449; Squal_synth.
DR Pfam; PF00494; SOS_PSY; 1.
DR TIGRFAMs; TIGR01559; squal_synth; 1.
DR PROSITE; PS01044; SQUALEN_PHTOEN_SYN_1; 1.
DR PROSITE; PS01045; SQUALEN_PHTOEN_SYN_2; 1.
KW Transferase.
SQ SEQUENCE 413 AA; 47289 MW; 9AEC733BAE14B4F3 CRC64;

Query Match 81.5%; Score 1671; DB 10; Length 413;
Best Local Similarity 81.4%; Pred. No. 1.3e-130;
Matches 311; Conservative 34; Mismatches 35; Indels 2; Gaps 1;
Qy 1 MGSGLTMLRYDDIYPLLMKRAIEKAEKQIPPEPHWGCYSMLHKVSRFSFSLVIOQLNT 60
Db 1 MGSGLGAIVRHDEVPYPLLMKRAARHAEKQIPPEPHWAFCTYMLLKVSRSFALVIOQLDP 60
Qy 61 ELRNVCVFVLVRALDTVEDDTSIPTDEKVPILIAFHRHIYDTDWHYSCGTKEYKILMD 120
Db 61 QLNRNVCIFYLVRALDTVEDDTSIATDVKEPILIAFHRHIYDRDWHFSCGTKEYKVLMD 120
Qy 121 QFHVHVAFLKGLKGYQAEIEITRMGAGMAKFCOEVEVDYDDEYCHYVAGLVGLGL 180
Db 121 QFHVHVAFLKGLKGYQAEIEITRMGAGMAKFCINEVEVDYDDEYCHYVAGLVGLGL 180
Qy 181 SKLFLAAGSEVLTPDWEAISNSMGLFLOKTNIIIRDYLEIDINEIPKSRMFPRWIKGYAD 240
Db 181 SKLFLAAGSEVLTPDWEAISNSMGLFLOKTNIIIRDYLEIDINEIPKSRMFPRWIKGYAD 240
Qy 241 KLEDLKYEENTNKSVOCLNEMVTNALMHIEDCLKYMVSRLDPSIFRFAIPQIMAGTGLA 300
Db 239 KLEDLKYEENSVAQVCLNDMVTNALLHABDCLKYSALRDMISIFRFAIPQIMAGTGLA 298
Qy 301 LCYNNEQVFRGVVGLRGLTAKVIDRTKTMADYVGFYDFSCMLTKYKVDKNDPNASKTLN 360
Db 299 LCYNNEQVFRGVVGLRGLTAKVIDRTKTMADYVGFYDFSCMLTKYKVDKNDPNASKTLN 358

Qy 361 RLEAVOKLCRDAGVLQNRKSYV 382
Db 359 RLEAIQKTCRESGLLSKRKPYI 380
RESULT 5
Q48666 PRELIMINARY; PRT; 415 AA.
AC Q48666;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Squalene synthase (EC 2.5.1.21).
OS Panax ginseng (Korean ginseng).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; campanulids; Apiales; Araliaceae; Panax.
OX NCBI_TaxID=4054;
RN [1]
RP SEQUENCE FROM N.A.
RA Suzuki H.;
RT "Isolation and characterization of a cDNA encoding the squalene
RL synthase from Panax ginseng.";
DR EMBL; AB010148; BAA24289.1; -;
DR InterPro; IPR002060; Squ/phyt_synthase.
DR InterPro; IPR006449; Squal_synth.
DR Pfam; PF00494; SOS_PSY; 1.
DR TIGRFAMs; TIGR01559; squal_synth; 1.
DR PROSITE; PS01044; SQUALEN_PHTOEN_SYN_1; 1.
DR PROSITE; PS01045; SQUALEN_PHTOEN_SYN_2; 1.
KW Transferase.
SQ SEQUENCE 415 AA; 47055 MW; 2576E2F0DDCE51F7 CRC64;

Query Match 81.1%; Score 1663; DB 10; Length 415;
Best Local Similarity 79.8%; Pred. No. 6.1e-130;
Matches 305; Conservative 42; Mismatches 33; Indels 2; Gaps 1;
Qy 1 MGSGLTMLRYDDIYPLLMKRAIEKAEKQIPPEPHWGCYSMLHKVSRFSFSLVIOQLNT 60
Db 1 MGSGLGAILKHPEDFYPLLMKFAARHAEKQIPPEPHWAFCTYMLLKVSRSFGLVIOQLGP 60
Qy 61 ELRNVCVFVLVRALDTVEDDTSIPTDEKVPILIAFHRHIYDTDWHYSCGTKEYKILMD 120
Db 61 QLNRDVCIFYLVRALDTVEDDTSIPTDEKVPILMAFHRHIYDKWHFSCGTKEYKVLMD 120
Qy 121 QFHVHVAFLKGLKGYQAEIEITRMGAGMAKFCOEVEVDYDDEYCHYVAGLVGLGL 180
Db 121 EPHVSNAPFELGSGYQAEIEDITRMGAGMAKFCICKEVETINDYDEYCHYVAGLVGLGL 180
Qy 181 SKLFLAAGSEVLTPDWEAISNSMGLFLOKTNIIIRDYLEIDINEIPKSRMFPRWIKGYAD 240
Db 181 SKLFLAAGSEVLTPDWEAISNSMGLFLOKTNIIIRDYLEIDINEIPKSRMFPRWIKGYAD 240
Qy 241 KLEDLKYEENTNKSVOCLNEMVTNALMHIEDCLKYMVSRLDPSIFRFAIPQIMAGTGLA 300
Db 239 KLEDLKYEENSVAQVCLNDMVTNALLHABDCLKYSALRDMISIFRFAIPQIMAGTGLA 298
Qy 301 LCYNNEQVFRGVVGLRGLTAKVIDRTKTMADYVGFYDFSCMLTKYKVDKNDPNASKTLN 360
Db 299 LCYNNEQVFRGVVGLRGLTAKVIDRTKTMADYVGFYDFSCMLTKYKVDKNDPNASKTLN 358
Qy 361 RLEAVOKLCRDAGVLQNRKSYV 382
Db 359 RLEAIQKTCRESGLLSKRKPYI 380
RESULT 6
Q42761 PRELIMINARY; PRT; 412 AA.
ID Q42761
AC Q42761;
DT 01-NOV-1996 (TReMBLrel. 01, Created)

| | | |
|--|--|--|
| DT | 01-NOV-1996 (T-EMBLrel. 01, Last sequence update) | |
| DT | 01-MAR-2003 (T-EMBLrel. 23, Last annotation update) | |
| DE | Squalene synthase [EC 2.5.1.21]. | |
| OS | Glycerhiza glabra. | |
| OC | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; | |
| OC | Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; | |
| OC | eurosid 1; Fabales; Fabaceae; Papilionoideae; Galeae; Glycyrrhiza. | |
| OC | NCBI_TaxID=49827; | |
| RN | [1] | |
| RP | SEQUENCE FROM N.A. | |
| RX | MEDLINE=97070595; PubMed=8913521; | |
| RA | Hayashi H., Hiraoka N., Ikeshiro Y.; | |
| RT | "Molecular cloning and functional expression of cDNAs for glycyrrhiza | |
| RT | glabra squalene synthase."; | |
| RL | Biol. Pharm. Bull. 19:1387-1389(1996). | |
| DR | EWBL; D86410; BAA13084.1; -. | |
| DR | InterPro; IPR002060; Squ_phyt_synthes. | |
| DR | InterPro; IPR008449; Squal_synth. | |
| DR | Pfam; PF00494; SQS_PSY; 1. | |
| DR | TIGRFAMs; TIGR01559; squal_synth; 1. | |
| DR | PROSITE; PS01044; SQUALEN_PHYTOEN_SYN_1; 1. | |
| DR | PROSITE; PS01045; SQUALEN_PHYTOEN_SYN_2; 1. | |
| KW | Transferase. | |
| SQ | SEQUENCE 412 AA; 47006 MW; 88D7EB9F9FD3595C CRC64; | |
| Query Match 80.9%; Score 1659; DB 10; Length 412; | | |
| Best Local Similarity 79.9%; Pred. No. 1.3e-129; | | |
| Matches 307; Conservative 42; Mismatches 33; Indels 2; Gaps 1; | | |
| QY | 1 MGSIGTMLRYPDIIYPLLMKRAIEKAEKQIPPEPHWGCYSMLHKVSRFSLVIQQLNT 60 | |
| DB | 1 MGSGLAILKHPDDIYPLKLTAAQAQAEKQIPREPHWSFCYTLHLHKVSRSPAFVIQQLGP 60 | |
| QY | 61 ELRNACVVFYLVLRALDVTDEDDTSTPTPEKYPILLAFHRIHYDTDWHYSCGTKEVKILMD 120 | |
| DB | 61 DLNRNACVIFYLVRALDVTDEDDTSTATDVKVPIILAFHRIHYDCDWHFSCGTKEVKILMD 120 | |
| QY | 121 OPHVYSAAFLEKGVQRAIEBITRRMGAGMAKFCQEVETVDDYDEYCHYVAGLVGLG 180 | |
| DB | 121 QPHVYSTAFLELCKYVQEAIEDITKRMGAGMAEFICEVETVDDYDEYCHYVAGLVGMGL 180 | |
| QY | 181 SKLFLAAGSEVLTDPWEALNSNGFLFLOKNTIIRDYLEIDINEIPKSRMFWPREIWKYAD 240 | |
| DB | 181 SKLFLAAGSGEDLASDY--LSNSMGLFLQKNTIIRDYLEIDINEIPKSRMFWPREIWKSYVN 238 | |
| QY | 241 KLEDLKYENTKNSVOCLNEMVTNALMHTEDCLKYMVSRLRDSIFRCAIPQIMAIGTLA 300 | |
| DB | 239 KLEDUKYENSVKAVOCLNDMVTNALMHAEDECLKYMAALRDLAIFRCAIPQIMAIGTLA 298 | |
| QY | 301 LCYNNEQVFRGVVVKLRGIGTAKVIDRTKTMADYGFYDFSCMLTKTKVDKNDPNASKTLN 360 | |
| DB | 299 LCYNNTIELFRGVVVMRGLTAKVIDRTKTMGDYGFDFPFASMLSKVDKNDPNATKTL 358 | |
| QY | 361 RLEAVQKLCEDAGVLQNRKSYND 384 | |
| DB | 359 RLEAIQKTCRESGLLNKRKSLLRN 382 | |
| RESULT 7 | | |
| Q9XJ31 PRELIMINARY; PRT; 411 AA. | | |
| ID | Q9XJ31 | |
| AC | Q9XJ31; | |
| DT | 01-NOV-1999 (T-EMBLrel. 12, Created) | |
| DT | 01-NOV-1999 (T-EMBLrel. 12, Last sequence update) | |
| DT | 01-MAR-2003 (T-EMBLrel. 23, Last annotation update) | |
| DE | Squalene synthase. | |
| GN | PSS1. | |
| OS | Solanum tuberosum (Potato). | |
| OC | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; | |
| OC | Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; | |
| OC | Asteridae; Lamids; Solanales; Solanaceae; Solanum. | |
| OX | NCBI_TaxID=4113; | |
| RN | [1] | |

```
DR Pfam; PF00494; SQS_PSY; 1.
DR TIGRFAMs; TIGR01559; squal_synth; 1.
DR PROSITE; PS01044; SQUALEN_PHYTOEN_SYN_1; 1.
DR PROSITE; PS01045; SQUALEN_PHYTOEN_SYN_2; 1.
DR SQUALEN_PHYTOEN_SYN_2; 1.
SQ SEQUENCE 411 AA; 46953 MW; 411771202057014 CRC64;

Query Match
Best Local Similarity 79.8%; Score 1637; DB 10; Length 411;
Matches 299; Conservative 44; Mismatches 37; Indels 2; Gaps 1;

QY 1 MGSIGTMLRYPDDIYPLLLKMKRAIEAEKQIPPEPHWGFCYSMLHKVSRFSLSVIOQLNT 60
Db 1 MGSRLAILKNPDDIYPLLLKMKRAIEAEKQIPPEPHWGFCYSMLHKVSRFSLSVIOQLPV 60
QY 61 ELRNACVCFVLRLALDTVEDDTSIPTDEKVPILIAFHRIHYDTDWHYSCGTKEYKILMD 120
Db 61 ELRDVAVCIFVLRLALDTVEDDTSIPTDEKVPILIDFHRHIYDNDWHFGCGTKEYKVLMD 120
QY 121 QPHHVSAAFLLEKGYQEAIEETTRMGAGMAKFIQCVETVDDYDEYCHYVAGLVGLGL 180
Db 121 QPHHVSAAFLLEKGYQEAIEETTRMGAGMAKFIQCVETVDDYDEYCHYVAGLVGLGL 180
QY 181 SKFLAAGSEVLTPDWEAISNSMGLFLOKTNIRDYLEDINEIPKSRMFMPREIWKYAD 240
Db 181 SKFLAAGSEVLTPDWEAISNSMGLFLOKTNIRDYLEDINEIPKSRMFMPREIWKYAD 240
QY 241 KLEDLKYENTNKSVOCLNEMVTNMLMHIEDCLKYMWSLRDPISIFRCAIPQIMAGITLA 300
Db 241 KLEDLKYENTNKSVOCLNEMVTNMLMHIEDCLKYMWSLRDPISIFRCAIPQIMAGITLA 300
QY 299 MCYDNEIEFRGVKMRGLTAKVIDRTKTMADYVGFYDFSCMLKTKVDKNDPNASKTLN 360
Db 299 MCYDNEIEFRGVKMRGLTAKVIDRTKTMADYVGFYDFSCMLKTKVDKNDPNASKTLN 360
QY 361 RLEAVOKLCRDAGVLQNRKSYV--NDKG 386
Db 361 RLEAVOKLCRDAGVLQNRKSYV--NDKG 386
QY 359 RLEAIQKTCRESGLLTKRKYVLRNESHG 386
Db 359 RLEAIQKTCRESGLLTKRKYVLRNESHG 386

RESULT 10
Q9XF02 PRELIMINARY; PRT; 411 AA.
AC Q9XF02;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Squalene synthase.
GN SS.
OS Capsicum annuum (Bell pepper).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; Lamiales; Solanales; Solanaceae; Capsicum.
OX NCBI_TaxID=4072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. subicho; TISSUE=leaf;
RA Yoon Y.H., Kim D.U., Shin D.H.;
RT "Cloning and characterization of Capsicum annuum squalene synthase and
RT regulation in response to UV-irradiation."
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF124842; AAD20626.1; -.
DR InterPro; IPR002060; Squ/phyt_synthase.
DR InterPro; IPR006449; Squal_synth.
DR Pfam; PF00494; SQS_PSY; 1.
DR TIGRFAMs; TIGR01559; squal_synth; 1.
DR PROSITE; PS01044; SQUALEN_PHYTOEN_SYN_1; 1.
DR PROSITE; PS01045; SQUALEN_PHYTOEN_SYN_2; 1.
DR SQUALEN_PHYTOEN_SYN_2; 1.
SQ SEQUENCE 411 AA; 47033 MW; 8D5B8EC50B31702D CRC64;

Query Match
Best Local Similarity 79.4%; Score 1629.5; DB 10; Length 411;
Matches 302; Conservative 42; Mismatches 41; Indels 3; Gaps 2;

QY 1 MGSIGTMLRYPDDIYPLLLKMKRAIEAEKQIPPEPHWGFCYSMLHKVSRFSLSVIOQLNT 60
Db 1 MGSRLAILKNPDDIYPLLLKMKRAIEAEKQIPPEPHWGFCYSMLHKVSRFSLSVIOQLPV 60
QY 61 ELRNACVCFVLRLALDTVEDDTSIPTDEKVPILIAFHRIHYDTDWHYSCGTKEYKILMD 120
Db 61 ELRDVAVCIFVLRLALDTVEDDTSIPTDEKVPILISFHQHIYDREWHFSCGTKEYKVLMD 120
QY 121 QPHHVSAAFLLEKGYQEAIEETTRMGAGMAKFIQCVETVDDYDEYCHYVAGLVGLGL 180
Db 121 QPHHVSAAFLLEKGYQEAIEETTRMGAGMAKFIQCVETVDDYDEYCHYVAGLVGLGL 180
QY 181 SKFLAAGSEVLTPDWEAISNSMGLFLOKTNIRDYLEDINEIPKSRMFMPREIWKYAD 240
Db 181 SKFLAAGSEVLTPDWEAISNSMGLFLOKTNIRDYLEDINEIPKSRMFMPREIWKYAD 240

Query Match
Best Local Similarity 79.6%; Score 1632; DB 10; Length 413;
Matches 305; Conservative 38; Mismatches 41; Indels 4; Gaps 2;

QY 1 MGSIGTMLRYPDDIYPLLLKMKRAIEAEKQIPPEPHWGFCYSMLHKVSRFSLSVIOQLNT 60
Db 1 MGSRLAILKNPDDIYPLLLKMKRAIEAEKQIPPEPHWGFCYSMLHKVSRFSLSVIOQLPV 60
```

```
Db 181 SKLPHASGKEDLASD--SLNSMGLFLQKTNIIIRDYLEDINEVVKPMFWRPFIWISKYN 238
Qy 241 KLEDLKYEENTNSVQCLNEMVTNALMHIEDCLKYMVSLRDPISIFRCAIPQIMAGITLA 300
Db 239 KLEELKYEENSVAQCLNDMVTNALSHVEDCLIMNSLRDPAIFRCAIPQWMAIGTLA 298
Qy 301 LCYNNEQVFRGVKRLGLTAKVIDRTKTWADVGFYDFSCMLKTKVDKNDPNASKTLN 360
Db 299 MCYDNIEVFRGVKWRGLTAKAIDRTKTWADVGFYDFSCMLKSKVNNNDPNATKTLK 358
Qy 361 RLEAVQKLCRDAGVQLNRKSYVNDKGP 388
Db 359 KLEAILKTCRDSGTLNKRKSYV-IKSEP 385

RESULT 11
Q9FUK1
ID Q9FUK1 PRELIMINARY; PRT; 418 AA.
AC Q9FUK1;
DT 01-MAR-2001 (TremBLrel. 16, Created)
DT 01-OCT-2001 (TremBLrel. 18, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE Squalene synthase.
OS Artemisia annua (Sweet wormwood).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; campanulids; Asterales; Asteraceae; Asteroideae;
OC Anthemideae; Artemisia.
OC NCBI_TaxID=35608;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=001;
RA Liu Y., Ye H.C., Li G.F.;
RT "Isolation of a cDNA encoding squalene synthase from Artemisia
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=001;
RA Liu Y., Ye H.C., Li G.F.;
RT "Isolation of squalene synthase (ASS) gene from Artemisia annua.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF302464; AAG14896.2; -.
DR EMBL; AF405310; AAK92539.1; -.
DR InterPro; IPR002060; Squ/phyt_synthse.
DR Pfam; PF00494; SQS_PSY; 1.
DR TIGRFAMs; TIGR01559; squal_synth; 1.
DR PROSITE; PS01045; SQUALEN PHYTOEN SYN 2; 1.
SQ SEQUENCE 418 AA; 47804 MW; C8DB9B9F03AA3856 CRC64;

Query Match 78.3%; Score 1605; DB 10; Length 418;
Best Local Similarity 78.5%; Pred. No. 4.1e-125;
Matches 300; Conservative 42; Mismatches 38; Indels 2; Gaps 1;

Qy 1 MGSIGTMLRYPDIDYPLLKMKRAIEKAEKQIPEPHWGFCYSMLHKVSRFSFLVIQQLNT 60
Db 1 MSSLKAVLKHDPDFDYPLLKMKRAIEKAEKQIPEPHWGFCYSMLHKVSRFSFLVIQQLNP 60
Qy 61 ELRNAVCVFYLVLRALDVTEDDTSIPTDEKVPILIAFHRIHYDTHWYSCGTKEYKILMD 120
Db 61 QLRAVAVCFYLVLRALDVTEDDTSIAADIKVPILIAFHRIHYDTHWYSCGTKEYKILMD 120
Qy 121 QFHVSAFLEKGYQAEIEETRRMGAGMAKFIQEVETVDDYDEYCHYVAGLVGLGL 180
Db 121 QFHVSTAFLELKGQYQAEIGDITMRMGAGMAKFIQEVETVDDYDEYCHYVAGLVGLGL 180
Qy 181 SKFLAAGSEVLTPDWEAISNSMGLFLQKTNIIIRDYLEDINEIPKSRMFWPREIMGYAD 240
Db 181 SKLPHSGTILFSD--SISNSMGLFLQKTNIIIRDYLEDINEIPKSRMFWPREIRSKYN 238
Qy 241 KLEDLKYEENTNSVQCLNEMVTNALMHIEDCLKYMVSLRDPISIFRCAIPQIMAGITLA 300
```

```
Db 239 KLEDLKYEEDSEKAVQCLNDMVTNALIHIEDCLKYMSQLKDPALFRCAIPQIMAGITLA 298
Qy 301 LCYNNEQVFRGVKRLGLTAKVIDRTKTWADVGFYDFSCMLKTKVDKNDPNASKTLN 360
Db 299 LCYNNEQVFRGVKRLGLTAKVIDRTKTWADVGFYDFSCMLKSKVNDPNATQITIT 358
Qy 361 RLEAVQKLCRDAGVQLNRKSYV 382
Db 359 RLEAAQKICKDSGTLNKRKSYI 380

RESULT 12
Q22106
ID Q22106 PRELIMINARY; PRT; 401 AA.
AC Q22106;
DT 01-JAN-1998 (TremBLrel. 05, Created)
DT 01-JAN-1998 (TremBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE Squalene synthase.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OC NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H84; TISSUE=Aerial parts;
RA Hata S.;
RT "cDNA cloning of squalene synthase genes from mono- and dicotyledonous
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB007502; BAA22558.1; -.
DR InterPro; IPR002060; Squ/phyt_synthse.
DR InterPro; IPR006449; Squal_synth.
DR Pfam; PF00494; SQS_PSY; 1.
DR TIGRFAMs; TIGR01559; squal_synth; 1.
DR PROSITE; PS01044; SQUALEN PHYTOEN SYN 1; 1.
DR PROSITE; PS01045; SQUALEN PHYTOEN SYN 2; 1.
SQ SEQUENCE 401 AA; 46040 MW; 1B6D3F89E77FED99 CRC64;

Query Match 70.5%; Score 1446; DB 10; Length 401;
Best Local Similarity 70.4%; Pred. No. 6.4e-112;
Matches 266; Conservative 57; Mismatches 49; Indels 6; Gaps 2;

Qy 1 MGSIGTMLRYPDIDYPLLKMKRAIEKAEKQIPEPHWGFCYSMLHKVSRFSFLVIQQLNT 60
Db 1 MGALSR----PEEVLALVKLRVAAGQIKRQIPPEEHWAFAYSMLQKVSRSFALVIQQLGP 56
Qy 61 ELRNAVCVFYLVLRALDVTEDDTSIPTDEKVPILIAFHRIHYDTHWYSCGTKEYKILMD 120
Db 57 ELRNAVCIFYLVLRALDVTEDDTSIPTEVKVPILQEFYRHIYNRDWHYSCGTNHYKMLMD 116
Qy 121 QFHVSAFLEKGYQAEIEETRRMGAGMAKFIQEVETVDDYDEYCHYVAGLVGLGL 180
Db 117 KFRVSTAFLELGEYQAEIEETRRMGAGMAKFIQEVETVDDYDEYCHYVAGLVGYGL 176
Qy 181 SKFLAAGSEVLTPDWEAISNSMGLFLQKTNIIIRDYLEDINEIPKSRMFWPREIMGYAD 240
Db 177 SRLFYAAGTEDALD--SLNSMGLFLQKTNIIIRDYLEDINEIPKSRMFWPREIWSKYAD 234
Qy 241 KLEDLKYEENTNSVQCLNEMVTNALMHIEDCLKYMVSLRDPISIFRCAIPQIMAGITLA 300
Db 235 KLEDFKYEENS KKAQVCLNNVTDALIHAECLQYNSALKDPALFRCAIPQIMAGITCA 294
Qy 301 LCYNNEQVFRGVKRLGLTAKVIDRTKTWADVGFYDFSCMLKTKVDKNDPNASKTLN 360
Db 295 LCYNNEQVFRGVKWRGLTAKVIDRTKTWADVGFYDFSCMLKSKVNDPNATQRTOR 354
Qy 361 RLEAVQKLCRDAGVQLNR 378
Db 355 RVDSIKQTCCKSSGLVKOR 372
```


RESULT 13

Q40472 PRELIMINARY; PRT; 376 AA.
 AC Q40472;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Squalene synthase.
 GS Nicotiana tabacum (Common tobacco).
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; lamids; Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98110157; PubMed=9448707;
 RA Davarenne T.P., Shin D.H., Back K., Yin S., Chappell J.;
 RT "Molecular characterization of tobacco squalene synthase and
 regulation in response to fungal elicitor";
 RL Arch. Biochem. Biophys. 349:205-215(1998).
 DR EMBL; U59683; AAB02945.1; -;
 DR InterPro; IPR002060; Squ/phyt_synthase.
 DR TIGRFAMs; TIGR01559; squal_synth; 1.
 DR Pfam; PF00494; SQS_PSY; 1.
 DR PROSITE; PS01044; SQUALEN_PHTOEN_SYN_1; 1.
 DR PROSITE; PS01044; SQUALEN_PHTOEN_SYN_2; 1.
 SQ SEQUENCE 376 AA; 42543 MW; 738253E680CC823D CRC64;

Query Match 70.1%; Score 1437.5; DB 10; Length 376;
 Best Local Similarity 70.4%; Pred. No. 3e-111;
 Matches 269; Conservative 42; Mismatches 34; Indels 37; Gaps 2;

QY 1 MGSILGTLRLYRDDIYPLLLKMKRAEAEKQIIPPEPHWGFCYMLHKVRSFSLVIOQLNT 60
 DB 1 MGSRLAILKMPDDIYPLVKKLAAHAEKQIIPPEPHWGFCYMLHKVRSFSLVIOQLPV 60
 QY 61 ELRNVCVFLVLRALDTVEDDTSIPTDEKVPILIAFHRIYDTHYSCGTKEYKILMD 120
 DB 61 ELRDVAVCFVLRALDTVEDDTSIPTDVKVPILISPHQVDRWHFSCGTKEYKILMD 120
 QY 121 QFHVSAFALEKGYQEAIEETIRRMAGMAKFIQCOEVETVDDYDEYHYVAGLVGLGL 180
 DB 121 QFHVSTAFLELRKHQQAEIETIRRMAGMAKFIQCOEVETVDDYDEYHYVAGLVGLGL 180
 QY 181 SKLFLAAGSVLPDWEAISNSMGLFLQKTIIRDYLEDINEIPKSRMFWPREWKGKAD 240
 DB 181 SKLFLHAGSGKEDLASD--SLSNSMGLFLQ----- 206
 QY 241 KLEDKYENTNKSVOCLNEMVTNALMHIEDCLKYMVSLRDPSTFRFCAIPQIMAGTGLA 300
 DB 207 --ELKYEDNSAKAVOCLNDMVTNALSHVEDCLTYNSALRDPSTFRFCAIPQVMAIGTGLA 263
 QY 301 LCYNNEQVFGVVKLRGLTAKVIDRTKTMADVYGAFYDFSCMLKTKVDKNDPNASKTLN 360
 DB 264 MCDYDNEVFGVVKMRGLTAKVIDRTIADYVGAFFDFSCMLKSKVNNNDPNATKTLK 323
 QY 361 RLEAVOKLCRDAGVLRNRSYV 382
 DB 324 RLEAILKTRDSGLTNKRKSYI 345

RESULT 14

O22105 PRELIMINARY; PRT; 403 AA.
 AC O22105;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Squalene synthase.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Ehrhartoideae; Oryzae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare; TISSUE=Root;
 RA Hata S.;
 RT "cDNA cloning of squalene synthase genes from mono- and dicotyledonous
 plants, and expression of the gene in rice";
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB007501; BAA23557.1; -;
 DR Gramene; O22105; -;
 DR InterPro; IPR002060; Squ/phyt_synthase.
 DR TIGRFAMs; TIGR01559; squal_synth; 1.
 DR Pfam; PF00494; SQS_PSY; 1.
 DR PROSITE; PS01044; SQUALEN_PHTOEN_SYN_1; 1.
 DR PROSITE; PS01044; SQUALEN_PHTOEN_SYN_2; 1.
 SQ SEQUENCE 403 AA; 46099 MW; F271C6EAI5403436 CRC64;

Query Match 69.5%; Score 1426; DB 10; Length 403;
 Best Local Similarity 69.5%; Pred. No. 3e-110;
 Matches 267; Conservative 56; Mismatches 57; Indels 4; Gaps 3;

QY 4 LGTMLRYDDIYPLLLKMKRAEAEKQIIPPEPHWGFCYMLHKVRSFSLVIOQLNTELR 63
 DB 1 MGVLSR-PEEVLPVLRVAAGRIKQIIPPEHWAFAITMLQVRSRFSALVIOQLGPDLR 59
 QY 64 NAVCVFLVLRALDTVEDDTSIPTDEKVPILIAFHRIYDTHYSCGTKEYKILMDQPH 123
 DB 60 NAVCIFYLVRALDTVEDDTSIPAAVKVPILKEFHRHRYNRDWHYSCGTKEYKILMDKFR 119
 QY 124 HVSAAFLKGYQEAIEETIRRMAGMAKFIQCOEVETVDDYDEYHYVAGLVGLGLSKL 183
 DB 120 LVSTAFLKGGQYQEAIEETIRLMGAGMAKFIQCOEVETVDDYDEYHYVAGLVGLSRL 179
 QY 184 FLAAGSEVLPDWEAISNSMGLFLQKTIIRDYLEDINEIPKSRMFWPREWKGKADKLE 243
 DB 180 FHAGGTEDLASD--SLSNSMGLFLQKTIIRDYLEDINEIPKSRMFWPREWKGKADKLE 237
 QY 244 DLKYENTNKSVOCLNEMVTNALMHIEDCLKYMVSLRDPSTFRFCAIPQIMAGTGLACY 303
 DB 238 DLKYENSEKAVOCLNDMVTNALSHVEDCLQYNSALRDPSTFRFCAIPQIMAGTGLACY 297
 QY 304 NNEQVFGVVKLRGLTAKVIDRTKTMADVYGAFYDFSCMLKTKVDKNDPNASKTLNRL 363
 DB 298 NNWVFGVVKMRGLTARVIDETNTMSDVYAFYEFSSIESKIDNNDPNASITRKRVD 357
 QY 364 AVOKLCRDAGVLRNRSYVNDKQ 387
 DB 358 AIKRTCKSSCLK-RRGYDLEKSK 380

RESULT 15

Q8LPN2 PRELIMINARY; PRT; 290 AA.
 ID Q8LPN2
 AC Q8LPN2;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Squalene synthase (Fragment).
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; lamids; Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. KY14;
 RA Davarenne T.P., Gosh A., Chappell J.;
 RT "Regulation of Squalene Synthase, a Key Enzyme in Sterol Biosynthesis
 in Nicotiana tabacum";
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY096801; AAM27472.1; -;

Thu Nov 13 10:01:48 2003

| | |
|----|--|
| DR | InterPro; IPRO02060; Squ/phyt_synthese. |
| DR | InterPro; IPRO06449; Squal_synth. |
| DR | Pfam; PF00494; SOS_PSY; 1. |
| DR | TIGRFAMs; TIGR01559; squal_synth; 1. |
| DR | PROSITE; PS01045; SQUALEN_PHYTOEN_SYN_2; 1. |
| FT | NON_TER 290 |
| SQ | SEQUENCE 290 AA; 33482 MW; 06CCA033E7F010EA CRC64; |
| | Query Match 61.3%; Score 1258; DB 10; Length 290; |
| | Best Local Similarity 78.8%; Pred. No. 1.8e-96; |
| | Matches 230; Conservative 30; Mismatches 30; Indels 2; Gaps 1; |
| Qy | 1 MGSLGTLRLYRDDIYPLLKKRALEKAEKOIPPEPHWGCYSMLHKVSRFSFLVIQQLT 60 |
| Dd | 1 MGSLRAILKNPEDLLPYLVLKLARHAEKOIPSPHWGCYSMLHKVSRFSALVIQLPV 60 |
| Qy | 61 EURNAVCVFYLVLRALDVTVEDDTSIPTDEKVPILIAFHRHIYDTDMHWSCGTKEYKIUMD 120 |
| Dd | 61 ELRDVCIFYLVLRALDVTVEDDTSIPTDVKVPLISFHQHVVDREMHFSCGTKEYKVLM D 120 |
| Qy | 121 QPHHSVAAPLEKEKYQBAIEITRRWGAGNAKFICOEVTDDDYCHYVAGIVGLGL 180 |
| Dd | 121 QPHHVSHTAFLELRHYOQAIEDITMRWGAGNAKFICEVEETDDDYCHYEAGIVGLGL 180 |
| Qy | 181 SKLFLAAGSEVLTDPWEAISNSMGLFLOKTNIIRDYLELINEIPKRMFWPREIWGYAD 240 |
| Dd | 181 SKLFHASGHEDLASD--SLSNSMGLFLOKTNIIRDYLELINEPKRCMFWPRESWKYVN 238 |
| Qy | 241 KLEDLKYEENTNKSVOCLNEWVTNALMHIEDCLKYKWSLRDPSIFFRCAIQ 292 |
| Dd | 239 KLEELKYENSAAVQCCLNDMVNTALSHEVDCLTYNSALRDPSIFFRCAIQ 290 |

Search completed: November 12, 2003, 16:30:26
Job time : 43 secs